PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C07K 14/00	A2	(11) International Publication Number:	WO 00/26245
		(43) International Publication Date:	11 May 2000 (11.05.00)

(21) International Application Number: PCT/US99/26048

(22) International Filing Date: 4 November 1999 (04.11.99)

(30) Priority Data:

60/172,255	4 November 1998 (04.11.98)	US
60/172,252	24 November 1998 (24.11.98)	US
60/172,214	22 December 1998 (22.12.98)	US
60/121,896	26 February 1999 (26.02.99)	US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications

US	60/172,255 (CIP)
Filed on	4 November 1998 (04.11.98)
US	60/172,252 (CIP)
Filed on	24 November 1998 (24.11.98)
US	60/172,214 (CIP)
Filed on	22 December 1998 (22.12.98)
US	60/121,896 (CIP)
Filed on	26 February 1999 (26.02.99)

(71) Applicant (*for all designated States except US*): INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): HILLMAN, Jennifer, L. [US/US]; 230 Monroe Drive, #12, Mountain View, CA 94040 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). CORLEY, Neil, C. [US/US]; 1240 Dale Avenue, #30, Mountain View, CA 94040 (US). GUEGLER, Karl, J. [CH/US]; 1048 Oakland Avenue, Menlo Park, CA 94025 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). AZIMZAI, Yalda [US/US]; 2045 Rock Springs Drive, Hayward, CA 94545 (US). LU, Dyung, Aina, M. [US/US]; 55 Park Belmont Place, San Jose, CA 95136 (US).

(74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: MEMBRANE TRANSPORT PROTEINS

```
1   M - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  1720440
1   M R A P G A G T A S V A S L A L L W F L G L P W T W S A A A  GI 2612939

2   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  1720440
31  A F C V Y V G G G G W R F L R I V C K T A R R D L F G L S V  GI 2612939

2   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  1720440
61  L I R V R L E L R R H R R A G D T I P C I F Q A V A R R Q P  GI 2612939

2   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  1720440
91  E R L A L V D A S S G I C W T F A Q L D T Y S N A V A N L F  GI 2612939

2   - - - - - - - - - - - - - - - E N R N E F V G L W L G M A K  1720440
121 R Q L G F A P G D V V A V F L E G R P E F V G L W L G L A K  GI 2612939

17  L G V E A A L I N T N L R R D A L L H C L T T S R A R A L V  1720440
151 A G V V A A L L N V N L R R E P L A F C L G T S A A K A L I  GI 2612939
```

(57) Abstract

The invention provides human membrane transport proteins (MTRP) and polynucleotides which identify and encode MTRP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of MTRP.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece	**ML**	Mali	**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**MN**	Mongolia	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MR**	Mauritania	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MW**	Malawi	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MX**	Mexico	**US**	United States of America
CA	Canada	**IT**	Italy	**NE**	Niger	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NL**	Netherlands	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NO**	Norway	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NZ**	New Zealand	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**PL**	Poland		
CM	Cameroon	**KR**	Republic of Korea	**PT**	Portugal		
CN	China	**KZ**	Kazakstan	**RO**	Romania		
CU	Cuba	**LC**	Saint Lucia	**RU**	Russian Federation		
CZ	Czech Republic	**LI**	Liechtenstein	**SD**	Sudan		
DE	Germany	**LK**	Sri Lanka	**SE**	Sweden		
DK	Denmark	**LR**	Liberia	**SG**	Singapore		
EE	Estonia						

MEMBRANE TRANSPORT PROTEINS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of membrane transport proteins and to the use of these sequences in the diagnosis, treatment, and prevention of membrane transport disorders; immune/inflammatory disorders; and cell proliferative disorders including cancer.

BACKGROUND OF THE INVENTION

Eukaryotic cells are bound by a lipid bilayer membrane and subdivided into functionally distinct, membrane bound compartments. The membranes maintain essential differences between the cytosol, the extracellular environment, and the contents of intracellular organelles such as the Golgi or the endoplasmic reticulum. As lipid membranes are highly impermeable to most polar molecules, transport of essential nutrients; metal ions such as K^+, NH_4^+, P_i, SO_4^{2-}; sugars; vitamins; metabolic waste products; cell signaling molecules; drugs; peptides; and proteins and other macromolecules across lipid membranes and between organelles must be mediated by a variety of transport molecules. Many transport mechanisms are substrate specific, with each transport protein carrying particular members of a molecular class, such as ions, sugars, or amino acids, across membranes. For example, amino acids are imported into cells via specific amino acid permeases.

Transport proteins are multi-pass transmembrane proteins, which either actively transport molecules across the membrane or passively allow them to cross. Active transport involves directional pumping of a solute across the membrane, usually against an electrochemical gradient. Active transport is tightly coupled to a source of metabolic energy, such as ATP hydrolysis or an electrochemically favorable ion gradient. Passive transport involves the movement of a solute down its electrochemical gradient. Transport proteins can be further classified as either carrier proteins or channel proteins. Carrier proteins, which can function in active or passive transport, bind to a specific solute to be transported and undergo a conformational change which transfers the bound solute across the membrane. Channel proteins, which only function in passive transport, form hydrophilic pores across the membrane. When the pores open, specific solutes, such as inorganic ions, pass through the membrane and down the electrochemical gradient of the solute.

Transport proteins play roles in antibiotic resistance, toxin secretion, ion balance, synaptic neurotransmission, kidney function, intestinal absorption, tumor growth, and other diverse cell functions (Griffith, J. and C. Sansom (1998) The Transporter Facts Book, Academic Press, San Diego CA, pp. 3-29). A variety of human inherited diseases are caused by mutation of transport proteins. For example, cystinuria is an inherited disease that results from the inability to transport cystine, the

disulfide-linked dimer of cysteine, from the urine into the blood. Accumulation of cystine in the urine leads to the formation of cystine stones in the kidneys. Also, many transport proteins are composed of subunits that may confer specificity for the tissue in which the transport mechanism functions, and are therefore associated with tissue-specific disorders. Examples of transport proteins include facilitative transporters, the secondary active symporters and antiporters driven by ion gradients, and active ATP binding cassette transporters involved in multiple-drug resistance and targeting of antigenic peptides to MHC Class I molecules, and the E1-E2 cation transport ATPases.

Carrier proteins which transport a single solute from one side of the membrane to the other are called uniporters. In contrast, coupled transporters link the transfer of one solute with simultaneous or sequential transfer of a second solute, either in the same direction (symport) or in the opposite direction (antiport). For example, intestinal and kidney epithelium contains a variety of symporter systems wherein the movement of sodium into the cell down its electrochemical gradient co-transports a second solute into the cell. The sodium gradient that provides the driving force for solute uptake is maintained by the ubiquitous Na^+/K^+ ATPase. Sodium-coupled transporters include the mammalian glucose transporter (SGLT1), iodide transporter (NIS), and multivitamin transporter (SMVT). These three transporters have twelve putative transmembrane segments, extracellular glycosylation sites, and cytoplasmically-oriented N- and C-termini. NIS plays a crucial role in the evaluation, diagnosis, and treatment of various thyroid pathologies because it is the molecular basis for radioiodide thyroid-imaging techniques and for specific targeting of radioisotopes to the thyroid gland (Levy, O. et al. (1997) Proc. Natl. Acad. Sci. USA 94:5568-5573). SMVT is expressed in the intestinal mucosa, kidney, and placenta, and is implicated in the transport of the water-soluble vitamins, e.g., biotin and pantothenate (Prasad, P.D. et al. (1998) J. Biol. Chem. 273:7501-7506).

The largest and most diverse family of transport proteins is the ATP-binding cassette (ABC) transporters. As a family, ABC transporters can transport substances that differ markedly in chemical structure and size, ranging from small molecules such as ions, sugars, amino acids, peptides, and phospholipids, to lipopeptides, large proteins, and complex hydrophobic drugs. Each ABC transporter consists of four modules: two nucleotide-binding domains (NBDs), which hydrolyze ATP to supply the energy required for transport; and two membrane-spanning domains (MSDs), which may form membrane channels. The NBDs consist of approximately two hundred conserved amino acid residues while the MSDs each contain six putative transmembrane segments. (See, e.g., Saurin, W. et al. (1994) Mol. Microbiol. 12:993-1004; Shani, N. et al. (1996) J. Biol. Chem. 271:8725-8730; Koster, W. and B. Bohm (1992) Mol. Gen. Genet. 232:399-407.) The four ABC transporter modules may be encoded by a single gene, as is the case for the cystic fibrosis transmembrane conductance regulator (CFTR), or by separate genes. When encoded by separate genes, each gene product

contains a single NBD and MSD. These "half-molecules" form homo- and heterodimers, such as Tap1 and Tap2, the endoplasmic reticulum-based major histocompatibility (MHC) peptide transport system associated with antigen processing (Androlewicz, M.J. et al. (1994) Proc. Natl. Acad. Sci. USA 91:12716-12720).

Several genetic diseases are attributed to defects in ABC transporters, including the following diseases and their corresponding proteins: cystic fibrosis (CFTR, an ion channel; Welsh, M.J. and A.E. Smith (1993) Cell 73:1251-1254); X-linked adrenoleukodystrophy, an inborn error of peroxisomal β-oxidation of very long chain fatty acids (adrenoleukodystrophy protein, ALDP); Zellweger syndrome, an inborn error of peroxisome biogenesis (peroxisomal membrane protein-70, PMP70); and hyperinsulinemic hypoglycemia (sulfonylurea receptor, SUR). The ABC transporters known as P-glycoproteins, or multidrug resistance (MDR) proteins, are associated with resistance to a wide range of hydrophobic drugs (MDR1; Gottesman, M.M. and I. Pastan (1993) Annu. Rev. Biochem. 62:385-427) or with phosphatidylcholine transport (MDR2; Ruetz, S. and P. Gros (1994) Cell 77:1071-1081). MDR is common in cancer cells, and contributes to low efficacy or failure of chemotherapy (Taglight, D. and S. Michaelis (1998) Methods Enzymol. 292:131-163). MDR is mediated by transporters, e.g., P-glycoproteins or the multidrug resistance-associated protein MRP, that normally function in the liver, intestines, and kidney to move toxic substances from the cytosol into the bile, intestinal lumen, or urine. In cancerous cells, these transporters extrude chemotherapeutic agents into the extracellular space, thereby conferring drug resistance. Recently, an ABC transporter-type protein was isolated from a human leukemia cell line. This transporter, termed the anthracycline resistance associated protein (GI 1279457, SEQ ID NO:42), is overexpressed in a multidrug resistant leukemia cell sub-line, and has sequence homology with other multidrug-resistance associated proteins including MRP (Longhurst, T.J. et al. (1996) Br. J. Cancer 74:1331-1335).

Transport of fatty acids across the plasma membrane can occur by diffusion, a high capacity, low affinity process. However, under normal physiological conditions a significant fraction of fatty acid transport appears to occur via a high affinity, low capacity protein-mediated transport process. Fatty acid transport protein (FATP), an integral membrane protein with four transmembrane segments, is expressed in tissues exhibiting high levels of plasma membrane fatty acid flux, such as muscle, heart, and adipose. Expression of FATP is upregulated in 3T3-L1 cells during adipose conversion, and expression in COS7 fibroblasts elevates the cells' uptake of long-chain fatty acids. Expression studies suggest a role for FATP in lipid metabolism, obesity, and type II diabetes mellitus (Hui, T.Y. et al. (1998) J. Biol. Chem. 273:27420-27429).

E1-E2 (or P-type) ATPases constitute a superfamily of cation transporters present in both

prokaryotes and eukaryotes that mediate membrane flux of all biologically relevant cations. These ATPases are postulated to exist in two different conformational states, designated E1 and E2, during the course of the ATP hydrolysis reaction, and to conserve the energy from ATP hydrolysis in the form of an acyl phosphate, primarily an aspartyl phosphate. Members of this family are divided into four major groups; the Ca^{+2}-transporting ATPases, Na^{+}/K^{+}-and gastric H^{+}/K^{+}-transporting ATPases, plasma membrane H^{+}-transporting ATPases (proton pumps), and the bacterial P-type ATPases (BLOCKS: BL00154, P-type cation-transporting ATPase superfamily signature).

The metabolism of amino acids is complex and highly regulated. While cells are capable of creating most amino acids de novo, the import of amino acids into cells via specific amino acid permease proteins is vital for maintaining the appropriate and complete availability of all necessary amino acids. This is particularly important during cell proliferation and differentiation. In addition to their role as protein building blocks, amino acids also serve as precursors for a variety of other important macromolecules. For example, the hormone thyroxine, the pigment melanin, and the neurotransmitters histamine, epinephrine, and serotonin are produced from various amino acid precursors, including histidine, tyrosine, and tryptophan. A component of sphingolipid formation, sphingosine, is derived from serine. Porphyrin rings, which are components of heme molecules, use glycine as a nitrogen donor. Significant portions of the ring structures of purines and pyrimidines, components of nucleic acids, are formed from the breakdown of numerous amino acids. Amino acids are also important in energy metabolism. Unlike fatty acids and glucose, amino acids cannot be stored in the cell, so excess amino acids are fed into the citric acid cycle to produce energy molecules including fatty acids, ketone bodies, and glucose. Thus, precise control of amino acid metabolism is extremely important to both proliferating and non-proliferating cells.

The E16 gene, cloned from human peripheral blood lymphocytes, encodes a 241 amino acid integral membrane protein with multiple predicted transmembrane domains (Gaugitsch, H.W. et al. (1992) J. Biol. Chem. 267:11267-11273). E16 gene expression is closely linked to cellular activation and division. In myeloid and lymphoid cells, E16 transcripts are rapidly induced and rapidly degraded after stimulation. This pattern of expression resembles the kinetics seen for proto-oncogenes and lymphokines in the T cell system. Elevated levels of E16 expression were detected in colonic, gastric, and breast adenocarcinomas, and in lymphoma, while little or no E16 expression was detected in normal (non-cancerous) human tissues such as adult brain, lung, liver, colon, esophagus, stomach, or kidney, nor in four-month fetal brain, lung, liver, or kidney (Wolf, D.A. et al. (1996) Cancer Res. 56:5012-5022; Gaugitsch et al., supra). E16 was detected in every cell line tested. Its presence in rapidly dividing cell lines and its absence in human tissues with low proliferative potential suggest that E16 is directly involved in the cell division process, where it helps provide important

building blocks for energy metabolism, biochemical synthetic pathways, and protein synthesis.

Post-translational modification of polypeptides occurs in the lumen of the Golgi apparatus. Such modifications include, for example, the addition of sugar molecules by enzymes such as N-acetylglucosaminyltransferase, to produce glycoproteins. The sugar-donating molecules in this reaction are typically nucleotide sugars, such as uridine diphosphate-galactose (UDP-Gal). UPD-Gal and other nucleotide sugars are transported from the cytosol into the Golgi apparatus by specific transporter molecules. The availability of these nucleotide sugars can regulate which glycoproteins are synthesized, and therefore has a significant impact on cellular function (Toma, L. et al. (1996) J. Biol. Chem. 271:3897-3901; Guillen, E. et al. (1998) Proc. Natl. Acad. Sci. USA 95:7888-7892).

The discovery of new membrane transport proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of membrane transport disorders; immune/inflammatory disorders; and cell proliferative disorders including cancer.

SUMMARY OF THE INVENTION

The invention features substantially purified polypeptides, membrane transport proteins, referred to collectively as "MTRP" and individually as "MTRP-1," "MTRP-2," "MTRP-3," "MTRP-4," "MTRP-5," "MTRP-6," "MTRP-7," "MTRP-8," "MTRP-9," "MTRP-10," "MTRP-11," "MTRP-12," "MTRP-13," "MTRP-14," "MTRP-15," "MTRP-16," and "MTRP-17." In one aspect, the invention provides a substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-17 and fragments thereof. The invention also includes a polypeptide comprising an amino acid sequence that differs by one or more conservative amino acid substitutions from an amino acid sequence selected from the group consisting of SEQ ID NO:1-17.

The invention further provides a substantially purified variant having at least 90% amino acid identity to at least one of the amino acid sequences selected from the group consisting of SEQ ID NO:1-17 and fragments thereof. The invention also provides an isolated and purified polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-17 and fragments thereof. The invention also includes an isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-17 and fragments thereof.

Additionally, the invention provides an isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-17 and fragments thereof. The

invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide encoding the polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1-17 and fragments thereof.

The invention also provides a method for detecting a polynucleotide in a sample containing nucleic acids, the method comprising the steps of: (a) hybridizing the complement of the polynucleotide sequence to at least one of the polynucleotides of the sample, thereby forming a hybridization complex; and (b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of a polynucleotide in the sample. In one aspect, the method further comprises amplifying the polynucleotide prior to hybridization.

The invention also provides an isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:18-34 and fragments thereof. The invention further provides an isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide sequence selected from the group consisting of SEQ ID NO:18-34 and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:18-34 and fragments thereof.

The invention further provides an expression vector containing at least a fragment of the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-17. In another aspect, the expression vector is contained within a host cell.

The invention also provides a method for producing a polypeptide, the method comprising the steps of: (a) culturing the host cell containing an expression vector containing a polynucleotide of the invention under conditions suitable for the expression of the polypeptide; and (b) recovering the polypeptide from the host cell culture.

The invention also provides a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-17 and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention further includes a purified antibody which binds to a polypeptide selected from the group consisting of SEQ ID NO:1-17 and fragments thereof. The invention also provides a purified agonist and a purified antagonist to the polypeptide.

The invention also provides a method for treating or preventing a disorder associated with decreased expression or activity of MTRP, the method comprising administering to a subject in need of such treatment an effective amount of a pharmaceutical composition comprising a substantially

purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-17 and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention also provides a method for treating or preventing a disorder associated with increased expression or activity of MTRP, the method comprising administering to a subject in need of such treatment an effective amount of an antagonist of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-17 and fragments thereof.

BRIEF DESCRIPTION OF THE FIGURES AND TABLES

Figures 1A, 1B, 1C, and 1D show the amino acid sequence alignment between MTRP-3 (Incyte Clone ID 1720440; SEQ ID NO:3) and mouse fatty acid transport protein (GI 2612939; SEQ ID NO:35), produced using the multisequence alignment program of LASERGENE software (DNASTAR, Madison WI).

Figures 2A, 2B, 2C, and 2D show the amino acid sequence alignment between MTRP-4 (Incyte Clone ID 2274290; SEQ ID NO:4) and Schistosoma mansoni SMDR1 (GI 425474; SEQ ID NO:36), produced using the multisequence alignment program of LASERGENE software (DNASTAR).

Figures 3A, 3B, 3C, and 3D show the amino acid sequence alignment between MTRP-5 (Incyte Clone ID 2740029; SEQ ID NO:5) and rat sodium-dependent multivitamin transporter (GI 3015617; SEQ ID NO:37), produced using the multisequence alignment program of LASERGENE software (DNASTAR).

Table 1 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOs), clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding MTRP.

Table 2 shows features of each polypeptide sequence, including potential motifs, homologous sequences, and methods, algorithms, and searchable databases used for analysis of MTRP.

Table 3 shows the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis; diseases, disorders, or conditions associated with these tissues; and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones encoding MTRP were isolated.

Table 5 shows the tools, programs, and algorithms used to analyze MTRP, along with applicable descriptions, references, and threshold parameters.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"MTRP" refers to the amino acid sequences of substantially purified MTRP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of MTRP. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of MTRP either by directly interacting with MTRP or by acting on components of the biological pathway in which MTRP participates.

An "allelic variant" is an alternative form of the gene encoding MTRP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding MTRP include those sequences with deletions,

insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as MTRP or a polypeptide with at least one functional characteristic of MTRP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding MTRP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding MTRP. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent MTRP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of MTRP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of MTRP. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of MTRP either by directly interacting with MTRP or by acting on components of the biological pathway in which MTRP participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind MTRP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the

translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition containing a nucleic acid sequence which is complementary to the "sense" strand of a specific nucleic acid sequence. Antisense molecules may be produced by any method including synthesis or transcription. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form duplexes and to block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic MTRP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The terms "complementary" and "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" bonds to the complementary sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acid strands, and in the design and use of peptide nucleic acid (PNA) molecules.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding MTRP or fragments of MTRP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be

associated with a stabilizing agent such as a carbohydrate.. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, extended using the XL-PCR kit (Perkin-Elmer, Norwalk CT) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from the overlapping sequences of one or more Incyte Clones and, in some cases, one or more public domain ESTs, using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that, when made, least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

Original Residue	Conservative Substitution
Ala	Gly, Ser
Arg	His, Lys
Asn	Asp, Gln, His
Asp	Asn, Glu
Cys	Ala, Ser
Gln	Asn, Glu, His
Glu	Asp, Gln, His
Gly	Ala
His	Asn, Arg, Gln, Glu
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Ile
Phe	His, Met, Leu, Trp, Tyr
Ser	Cys, Thr
Thr	Ser, Val
Trp	Phe, Tyr
Tyr	His, Phe, Trp
Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to the chemical modification of a polypeptide sequence, or a polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "fragment" is a unique portion of MTRP or the polynucleotide encoding MTRP which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50% of a polypeptide) as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:18-34 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:18-34, for example, as distinct from any other sequence in the same genome. A fragment of SEQ ID NO:18-34 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:18-34 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:18-34 and the region of SEQ ID NO:18-34 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-17 is encoded by a fragment of SEQ ID NO:18-34. A fragment of SEQ ID NO:1-17 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-17. For example, a fragment of SEQ ID NO:1-17 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-17. The precise length of a fragment of SEQ ID NO:1-17 and the region of SEQ ID NO:1-17 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

The term "similarity" refers to a degree of complementarity. There may be partial similarity or complete similarity. The word "identity" may substitute for the word "similarity." A partially

complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as "substantially similar." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization, and the like) under conditions of reduced stringency. A substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under conditions of reduced stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as reduced stringency conditions require that the binding of two sequences to one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the substantially similar sequence or probe will not hybridize to the second non-complementary target sequence.

The phrases "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequence pairs.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at http://www.ncbi.nlm.nih.gov/BLAST/. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2

Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at http://www.ncbi.nlm.nih.gov/gorf/bl2.html. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62
Reward for match: 1
Penalty for mismatch: -2
Open Gap: 5 and Extension Gap: 2 penalties
Gap x drop-off: 50
Expect: 10
Word Size: 11
Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the hydrophobicity and acidity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of

polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for stable mitotic chromosome segregation and maintenance.

The term "humanized antibody" refers to antibody molecules in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of identity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and

may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Generally, such wash temperatures are selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions for nucleic acid hybridization are well known and can be found in Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect

cellular and systemic defense systems.

The term "microarray" refers to an arrangement of distinct polynucleotides on a substrate.

The terms "element" and "array element" in a microarray context, refer to hybridizable polynucleotides arranged on the surface of a substrate.

The term "modulate" refers to a change in the activity of MTRP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of MTRP.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Probe" refers to nucleic acid sequences encoding MTRP, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the

specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel et al.,1987, Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis et al., 1990, PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques

such as those described in Sambrook, supra. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be use to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding MTRP, or fragments thereof, or MTRP itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably about 75% free, and most preferably about 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

"Transformation" describes a process by which exogenous DNA enters and changes a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection,

electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0 9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95% or at least 98% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence identity over a certain defined length of one of the polypeptides.

THE INVENTION

The invention is based on the discovery of new human membrane transport proteins (MTRP), the polynucleotides encoding MTRP, and the use of these compositions for the diagnosis, treatment, or prevention of membrane transport disorders; immune/inflammatory disorders; and cell proliferative disorders including cancer.

Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding MTRP. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide

and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which nucleic acids encoding each MTRP were identified, and column 4 shows the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. The Incyte clones in column 5 were used to assemble the consensus nucleotide sequence of each MTRP and are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of each of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3 shows potential phosphorylation sites; column 4 shows potential glycosylation sites; column 5 shows the amino acid residues comprising signature sequences and motifs; column 6 shows homologous sequences as identified by BLAST analysis; and column 7 shows analytical methods and in some cases, searchable databases to which the analytical methods were applied. The methods of column 7 were used to characterize each polypeptide through sequence homology and protein motifs.

As shown in Figures 1A, 1B, 1C, and 1D, MTRP-3 has chemical and structural similarity with mouse fatty acid transport protein (FATP; GI 2612939; SEQ ID NO:35). In particular, MTRP-3 and FATP share 65% identity. As shown in Figures 2A, 2B, 2C, and 2D, MTRP-4 has chemical and structural similarity with Schistosoma mansoni ATP-binding cassette family protein, SMDR-1 (GI 425474; SEQ ID NO:36). In particular, MTRP-4 and SMDR-1 share 38% identity. As shown in Figures 3A, 3B, 3C, and 3D, MTRP-5 has chemical and structural similarity with rat sodium-dependent multivitamin transporter (SMVT; GI 3015617; SEQ ID NO:37). In particular, MTRP-5 and SMVT share 82% identity.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding MTRP. The first column of Table 3 lists the nucleotide SEQ ID NOs. Column 2 lists tissue categories which express MTRP as a fraction of total tissues expressing MTRP. Column 3 lists diseases, disorders, or conditions associated with those tissues expressing MTRP as a fraction of total tissues expressing MTRP. Column 4 lists the vectors used to subclone each cDNA library.

Of particular note are the expression patterns of SEQ ID NO:30 and SEQ ID NO:31. SEQ ID NO:30 is expressed in only five libraries, of which at least four (80%) are associated with cell proliferation and at least one (20%) with inflammation. Two (40%) of the five libraries are associated with cardiovascular tissue, and one each (20%) with gastrointestinal, nervous, and reproductive tissues. SEQ ID NO:31 is expressed in only four libraries, of which at least three (75%) are associated with cell proliferation and at least two (50%) with inflammation or the immune response.

Two (50%) of the four libraries are associated with hematopoietic/immune tissue, and one each (25%) with cardiovascular and reproductive tissues.

The following fragments of the nucleotide sequences encoding MTRP are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:18-34 and to distinguish between SEQ ID NO:18-34 and related polynucleotide sequences. The useful fragments include the fragment of SEQ ID NO:18 from about nucleotide 110 to about nucleotide 154; the fragment of SEQ ID NO:19 from about nucleotide 759 to about nucleotide 839; the fragment of SEQ ID NO:20 from about nucleotide 1531 to about nucleotide 1578; the fragment of SEQ ID NO:21 from about nucleotide 538 to about nucleotide 597; the fragment of SEQ ID NO:22 from about nucleotide 2241 to about nucleotide 2294; the fragment of SEQ ID NO:23 from about nucleotide 116 to about nucleotide 145; the fragment of SEQ ID NO:24 from about nucleotide 60 to about nucleotide 89; the fragment of SEQ ID NO:25 from about nucleotide 160 to about nucleotide 189; the fragment of SEQ ID NO:26 from about nucleotide 763 to about nucleotide 792; the fragment of SEQ ID NO:27 from about nucleotide 43 to about nucleotide 72; the fragment of SEQ ID NO:28 from about nucleotide 361 to about nucleotide 405; the fragment of SEQ ID NO:29 from about nucleotide 35 to about nucleotide 79; the fragment of SEQ ID NO:30 from about nucleotide 206 to about nucleotide 250; the fragment of SEQ ID NO:31 from about nucleotide 71 to about nucleotide 115; the fragment of SEQ ID NO:32 from about nucleotide 161 to about nucleotide 205; the fragment of SEQ ID NO:33 from about nucleotide 364 to about nucleotide 408; and the fragment of SEQ ID NO:34 from about nucleotide 18 to about nucleotide 62. The polypeptides encoded by the specified fragments of SEQ ID NO:20-30 and SEQ ID NO:32-34 are useful, for example, as immunogenic peptides.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding MTRP were isolated. Column 1 references the nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

The invention also encompasses MTRP variants. A preferred MTRP variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the MTRP amino acid sequence, and which contains at least one functional or structural characteristic of MTRP.

The invention also encompasses polynucleotides which encode MTRP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:18-34, which encodes MTRP.

The invention also encompasses a variant of a polynucleotide sequence encoding MTRP. In particular, such a variant polynucleotide sequence will have at least about 75%, or alternatively at

least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding MTRP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:18-34 which has at least about 75%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:18-34. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of MTRP.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding MTRP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring MTRP, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode MTRP and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring MTRP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding MTRP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding MTRP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode MTRP and MTRP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding MTRP or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:18-34 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol.

152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Perkin-Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Perkin-Elmer). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Perkin-Elmer), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding MTRP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National

Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin-Elmer), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode MTRP may be cloned in recombinant DNA molecules that direct expression of MTRP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express MTRP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter MTRP-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

In another embodiment, sequences encoding MTRP may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, MTRP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g.,

Robergc, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Perkin-Elmer). Additionally, the amino acid sequence of MTRP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY.)

In order to express a biologically active MTRP, the nucleotide sequences encoding MTRP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding MTRP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding MTRP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding MTRP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding MTRP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding MTRP. These include, but are not limited to, microorganisms such as bacteria transformed

with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding MTRP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding MTRP can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding MTRP into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of MTRP are needed, e.g. for the production of antibodies, vectors which direct high level expression of MTRP may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of MTRP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; and Scorer, C.A. et al. (1994) Bio/Technology 12:181-184.)

Plant systems may also be used for expression of MTRP. Transcription of sequences encoding MTRP may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases

where an adenovirus is used as an expression vector, sequences encoding MTRP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses MTRP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of MTRP in cell lines is preferred. For example, sequences encoding MTRP can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk*$^-$ and *apr*$^-$ cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), ß glucuronidase and its substrate ß-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to

quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding MTRP is inserted within a marker gene sequence, transformed cells containing sequences encoding MTRP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding MTRP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding MTRP and that express MTRP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of MTRP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on MTRP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding MTRP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding MTRP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for

ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding MTRP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode MTRP may be designed to contain signal sequences which direct secretion of MTRP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding MTRP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric MTRP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of MTRP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the MTRP encoding sequence and the heterologous protein sequence, so that MTRP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of

fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled MTRP may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ^{35}S-methionine.

Fragments of MTRP may be produced not only by recombinant means, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, supra, pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for example, using the ABI 431A peptide synthesizer (Perkin-Elmer). Various fragments of MTRP may be synthesized separately and then combined to produce the full length molecule.

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of MTRP and membrane transport proteins, including amino acid transporters, ABC transporters, nucleotide-sugar transporters, transmembrane carrier proteins, and ATP-dependent transporter proteins. In addition, the expression of MTRP is closely associated with nervous, reproductive, and gastrointestinal tissues; cancer and other cell proliferative conditions; and with inflammation and the immune response. Therefore, MTRP appears to play a role in membrane transport disorders; immune/inflammatory disorders; and cell proliferative disorders including cancer. In the treatment of disorders associated with increased MTRP expression or activity, it is desirable to decrease the expression or activity of MTRP. In the treatment of disorders associated with decreased MTRP expression or activity, it is desirable to increase the expression or activity of MTRP.

Therefore, in one embodiment, MTRP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of MTRP. Examples of such disorders include, but are not limited to, a membrane transport disorder such as cystinuria, dibasicaminoaciduria, hypercystinuria, lysinuria, hartnup disease, tryptophan malabsorption, methionine malabsorption, histidinuria, iminoglycinuria, dicarboxylicaminoaciduria, cystinosis, renal glycosuria, glucose-galactose malabsorption, familial hypercholesterolemia, hypouricemia, familial hypophophatemic rickets, congenital chloridorrhea, cystic fibrosis, familial goiter, distal renal tubular acidosis, Menkes' disease, lethal diarrhea, nephrogenic diabetes insipidus, juvenile pernicious anemia, folate malabsorption, adrenoleukodystrophy, hereditary myoglobinuria, Zellweger syndrome, hyperinsulinemic hypoglycemia, akinesia, amyotrophic lateral sclerosis, ataxia telangiectasia, cystic fibrosis, Becker's muscular dystrophy, Bell's palsy, Charcot-Marie Tooth disease, diabetes mellitus, diabetes insipidus,

diabetic neuropathy, Duchenne muscular dystrophy, hyperkalemic periodic paralysis, normokalemic periodic paralysis, Parkinson's disease, malignant hyperthermia, multidrug resistance, myasthenia gravis, myotonic dystrophy, catatonia, tardive dyskinesia, dystonias, peripheral neuropathy, cerebral neoplasms, and prostate cancer; a cardiac disorder associated with transport such as angina, bradyarrythmia, tachyarrythmia, hypertension, Long QT syndrome, myocarditis, cardiomyopathy, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, thyrotoxic myopathy, ethanol myopathy, dermatomyositis, inclusion body myositis, infectious myositis, and polymyositis; a neurological disorder associated with transport such as Alzheimer's disease, amnesia, bipolar disorder, dementia, depression, epilepsy, Tourette's disorder, paranoid psychoses, and schizophrenia; and an other disorder associated with transport such as neurofibromatosis, postherpetic neuralgia, trigeminal neuropathy, sarcoidosis, sickle cell anemia, Wilson's disease, cataracts, infertility, pulmonary artery stenosis, sensorineural autosomal deafness, hyperglycemia, hypoglycemia, Grave's disease, goiter, Cushing's disease, Addison's disease, glucose-galactose malabsorption syndrome, and hypercholesterolemia; an immune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyenodocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation; a viral, bacterial, fungal, parasitic, protozoal, or helminthic infection; and trauma; a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia; and a cancer including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma; and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus.

In another embodiment, a vector capable of expressing MTRP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of MTRP including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified MTRP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of MTRP including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of MTRP may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of MTRP including, but not limited to, those listed above.

In a further embodiment, an antagonist of MTRP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of MTRP. Examples of such disorders include, but are not limited to, those membrane transport disorders; immune/inflammatory disorders; and cell proliferative disorders including cancer described above. In one aspect, an antibody which specifically binds MTRP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express MTRP.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding MTRP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of MTRP including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of MTRP may be produced using methods which are generally known in the art. In particular, purified MTRP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind MTRP. Antibodies to MTRP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with MTRP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to MTRP have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of MTRP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to MTRP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce MTRP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for MTRP may also be generated.

For example, such fragments include, but are not limited to, $F(ab')_2$ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between MTRP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering MTRP epitopes is generally used, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for MTRP. Affinity is expressed as an association constant, K_a, which is defined as the molar concentration of MTRP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple MTRP epitopes, represents the average affinity, or avidity, of the antibodies for MTRP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular MTRP epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the MTRP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of MTRP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington, DC; Liddell, J.E. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of MTRP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding MTRP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding MTRP may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding MTRP. Thus, complementary molecules or fragments may be used to modulate MTRP activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding MTRP.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding MTRP. (See, e.g., Sambrook, supra; Ausubel, 1995, supra.)

Genes encoding MTRP can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding MTRP. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector, and may last even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding MTRP. Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may be employed. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme

molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding MTRP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding MTRP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of MTRP, antibodies to MTRP, and mimetics, agonists, antagonists, or inhibitors of MTRP. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combining active compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to

characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents to increase the solubility of the compounds and allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acids. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of MTRP, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example MTRP or fragments thereof, antibodies of MTRP, and agonists, antagonists or inhibitors of MTRP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind MTRP may be used for the diagnosis of disorders characterized by expression of MTRP, or in assays to monitor patients being

treated with MTRP or agonists, antagonists, or inhibitors of MTRP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for MTRP include methods which utilize the antibody and a label to detect MTRP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring MTRP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of MTRP expression. Normal or standard values for MTRP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibody to MTRP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of MTRP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding MTRP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of MTRP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of MTRP, and to monitor regulation of MTRP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding MTRP or closely related molecules may be used to identify nucleic acid sequences which encode MTRP. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding MTRP, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the MTRP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:18-34 or from genomic sequences including promoters, enhancers, and introns of the MTRP gene.

Means for producing specific hybridization probes for DNAs encoding MTRP include the cloning of polynucleotide sequences encoding MTRP or MTRP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may

be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding MTRP may be used for the diagnosis of disorders associated with expression of MTRP. Examples of such disorders include, but are not limited to, a membrane transport disorder such as cystinuria, dibasicaminoaciduria, hypercystinuria, lysinuria, hartnup disease, tryptophan malabsorption, methionine malabsorption, histidinuria, iminoglycinuria, dicarboxylicaminoaciduria, cystinosis, renal glycosuria, glucose-galactose malabsorption, familial hypercholesterolemia, hypouricemia, familial hypophophatemic rickets, congenital chloridorrhea, cystic fibrosis, familial goiter, distal renal tubular acidosis, Menkes' disease, lethal diarrhea, nephrogenic diabetes insipidus, juvenile pernicious anemia, folate malabsorption, adrenoleukodystrophy, hereditary myoglobinuria, Zellweger syndrome, hyperinsulinemic hypoglycemia, akinesia, amyotrophic lateral sclerosis, ataxia telangiectasia, cystic fibrosis, Becker's muscular dystrophy, Bell's palsy, Charcot-Marie Tooth disease, diabetes mellitus, diabetes insipidus, diabetic neuropathy, Duchenne muscular dystrophy, hyperkalemic periodic paralysis, normokalemic periodic paralysis, Parkinson's disease, malignant hyperthermia, multidrug resistance, myasthenia gravis, myotonic dystrophy, catatonia, tardive dyskinesia, dystonias, peripheral neuropathy, cerebral neoplasms, and prostate cancer; a cardiac disorder associated with transport such as angina, bradyarrythmia, tachyarrythmia, hypertension, Long QT syndrome, myocarditis, cardiomyopathy, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, thyrotoxic myopathy, ethanol myopathy, dermatomyositis, inclusion body myositis, infectious myositis, and polymyositis; a neurological disorder associated with transport such as Alzheimer's disease, amnesia, bipolar disorder, dementia, depression, epilepsy, Tourette's disorder, paranoid psychoses, and schizophrenia; and an other disorder associated with transport such as neurofibromatosis, postherpetic neuralgia, trigeminal neuropathy, sarcoidosis, sickle cell anemia, Wilson's disease, cataracts, infertility, pulmonary artery stenosis, sensorineural autosomal deafness, hyperglycemia, hypoglycemia, Grave's disease, goiter, Cushing's disease, Addison's disease, glucose-galactose malabsorption syndrome, and hypercholesterolemia; an immune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyenodocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia

with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation; a viral, bacterial, fungal, parasitic, protozoal, or helminthic infection; and trauma; a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia; and a cancer including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma; and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus. The polynucleotide sequences encoding MTRP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered MTRP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding MTRP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding MTRP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding MTRP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of MTRP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding MTRP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from

normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding MTRP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding MTRP, or a fragment of a polynucleotide complementary to the polynucleotide encoding MTRP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

Methods which may also be used to quantify the expression of MTRP include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and

monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding MTRP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding MTRP on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, MTRP, its catalytic or immunogenic fragments, or

oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between MTRP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with MTRP, or fragments thereof, and washed. Bound MTRP is then detected by methods well known in the art. Purified MTRP can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding MTRP specifically compete with a test compound for binding MTRP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with MTRP.

In additional embodiments, the nucleotide sequences which encode MTRP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. [Attorney Docket No. PF-0633 P, filed November 4, 1998], U.S. Ser. No. [Attorney Docket No. PF-0645 P, filed November 24, 1998], U.S. Ser. No. [Attorney Docket No. PF-0657 P, filed December 22, 1998], and U.S. Ser. No. 60/121,896, are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some

tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), or pINCY (Incyte Pharmaceuticals, Palo Alto CA). Recombinant plasmids were transformed into competent E. coli cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5α, DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a

high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

cDNA sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Perkin-Elmer) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Perkin-Elmer) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example V.

The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 5 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments were generated using the default parameters specified by the clustal algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

The polynucleotide sequences were validated by removing vector, linker, and polyA

sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programing, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and PFAM to acquire annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene families. (See, e.g., Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:18-34. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

IV. Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel, 1995, supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in nucleotide databases such as GenBank or LIFESEQ (Incyte Pharmaceuticals). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\%\ \text{sequence identity} \times \%\ \text{maximum BLAST score}}{100}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and, with a product score of 70, the match will be exact. Similar molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding MTRP occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease/condition categories included cancer, inflammation, trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in Table 3.

V. Extension of MTRP Encoding Polynucleotides

The full length nucleic acid sequences of SEQ ID NO:18-34 were produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+}, $(NH_4)_2SO_4$, and β-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 µl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 µl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II

(Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μl to 10 μl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent E. coli cells. Transformed cells were selected on antibiotic-containing media, individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethysulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

In like manner, the nucleotide sequences of SEQ ID NO:18-34 are used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for such extension, and an appropriate genomic library.

VI. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:18-34 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a

SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

VII. Microarrays

A chemical coupling procedure and an ink jet device can be used to synthesize array elements on the surface of a substrate. (See, e.g., Baldeschweiler, supra.) An array analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the scanned images.

Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). Full-length cDNAs, ESTs, or fragments thereof corresponding to one of the nucleotide sequences of the present invention, or selected at random from a cDNA library relevant to the present invention, are arranged on an appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

VIII. Complementary Polynucleotides

Sequences complementary to the MTRP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring MTRP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same

procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of MTRP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the MTRP-encoding transcript.

IX. Expression of MTRP

Expression and purification of MTRP is achieved using bacterial or virus-based expression systems. For expression of MTRP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express MTRP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of MTRP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding MTRP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, MTRP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from MTRP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified MTRP obtained by these methods can be used directly in the

following activity assay.

X. Demonstration of MTRP Activity

ATPase activity associated with MTRP can be measured by hydrolysis of radiolabeled ATP-[γ-^{32}P], separation of the hydrolysis products by chromatographic methods, and quantitation of the recovered ^{32}P using a scintillation counter. The reaction mixture contains ATP-[γ-^{32}P] and varying amounts of MTRP in a suitable buffer incubated at 37°C for a suitable period of time. The reaction is terminated by acid precipitation with trichloroacetic acid and then neutralized with base, and an aliquot of the reaction mixture is subjected to membrane or filter paper-based chromatography to separate the reaction products. The amount of ^{32}P liberated is counted in a scintillation counter. The amount of radioactivity recovered is proportional to the ATPase activity of MTRP in the assay.

MTRP transport activity is assayed by measuring uptake of labeled substrates into Xenopus laevis oocytes. Oocytes at stages V and VI are injected with MTRP mRNA (10 ng per oocyte) and incubated for 3 days at 18°C in OR2 medium (82.5mM NaCl, 2.5 mM KCl, 1mM $CaCl_2$, 1mM $MgCl_2$, 1mM Na_2HPO_4, 5 mM Hepes, 3.8 mM NaOH , 50μg/ml gentamycin, pH 7.8) to allow expression of MTRP protein. Oocytes are then transferred to standard uptake medium (100mM NaCl, 2 mM KCl, 1mM $CaCl_2$, 1mM $MgCl_2$, 10 mM Hepes/Tris, pH 7.5). Uptake of various substrates (e.g., amino acids, sugars, drugs, and neurotransmitters) is initiated by adding a ^{3}H-labeled substrate to the oocytes. After 30 minutes of incubation, uptake is terminated by washing the oocytes three times in Na^+-free medium. Incorporation of ^{3}H is measured, and compared with controls. MTRP transport activity is proportional to the level of internalized ^{3}H-labeled substrate.

XI. Functional Assays

MTRP function is assessed by expressing the sequences encoding MTRP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT (Life Technologies) and pCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of

fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of MTRP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding MTRP and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding MTRP and other genes of interest can be analyzed by northern analysis or microarray techniques.

XII. Production of MTRP Specific Antibodies

MTRP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the MTRP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Perkin-Elmer) using fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-MTRP activity by, for example, binding the peptide or MTRP to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XIII. Purification of Naturally Occurring MTRP Using Specific Antibodies

Naturally occurring or recombinant MTRP is substantially purified by immunoaffinity

chromatography using antibodies specific for MTRP. An immunoaffinity column is constructed by covalently coupling anti-MTRP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing MTRP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of MTRP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/MTRP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and MTRP is collected.

XIV. Identification of Molecules Which Interact with MTRP

MTRP, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled MTRP, washed, and any wells with labeled MTRP complex are assayed. Data obtained using different concentrations of MTRP are used to calculate values for the number, affinity, and association of MTRP with the candidate molecules.

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	18	961344	BRSTTUT03	642329R1 (BRSTNOT03), 961344H1 (BRSTTUT03), 3149970H1 (ADRENON04), 3255621H1 (OVARTUN01), 4265773H1 (KIDNNOT32), 4641320H1 (PROSTMT03), 4875558H1 (COLDNOT01)
2	19	3128782	LUNGTUT12	140511X20R1 (FIBRNGT01), 580891H1 (BRAVTXT05), 1304328F1 (PLACNOT02), 1469890T6 (PANCTUT02), 1799816T6 (COLNNOT27), 1985910R6 (LUNGAST01), 2722244F6 (LUNGTUT10), 3128782F6 (LUNGTUT12), 3128782H1 (LUNGTUT12), 3276993F6 (PROSBPT06), SBMA03256F1
3	20	1720440	BLADNOT06	744485R1 (BRAITUT01), 859118R1 (BRAITUT03), 944049T1 (ADRENOT03), 1432755R1 (BEPINON01), 1720440H1 (BLADNOT06), 1808737X11C1 (PROSTUT12), 1812106F6 (PROSTUT12), 2192988F6 (THYRTUT03), 2192988X13F1 (THYRTUT03), 2192988X14F1 (THYRTUT03), 3384757H1 (ESOGNOT04)
4	21	2274290	PROSNON01	1732422F6 (BRSTTUT08), 2098563H1 (BRAITUT02), 2274290H1 (PROSNON01), 2274290X326D2 (PROSNON01), 2598580F6 (UTRSNOT10), 2779864F6 (OVARTUT03), 2864759F6 (KIDNNOT20), 2864759T6 (KIDNNOT20), 3221871H1 (COLNNON03)
5	22	2740029	BRSTTUT14	966363H1 (BRSTNOT05), 1000112R1 (BRSTNOT03), 1647057F6 (PROSTUT09), 2740029H1 (BRSTTUT14), 2740029X309F1 (BRSTTUT14), 2740029X321F1 (BRSTTUT14), 5068692H1 (PANCNOT23), 5193978H1 (LUNLTUT04), SBOA03895D1
6	23	2414415	HNT3AZT01	1508631H1 (LUNGNOT14), 1596418F1 (BRAINOT14), 2414415F6 (HNT3AZT01), 2414415H1 (HNT3AZT01), 2414415X300D1 (HNT3AZT01), 2902794H1 (DRGCNOT01), 3080373H1 (BRAIUNT01), 3554689H1 (SYNONOT01), 3881949F6 (SPLNNOT11), 4996983H1 (MYEPTXT02)
7	24	2466714	THYRNOT08	627910R6 (KIDNNOT05), 2358261R6 (LUNGFET05), 2358464F6 (LUNGFET05), 2466714H1 (THYRNOT08), 2726053F6 (OVARTUT05), 3845383H1 (DENDNOT01), 4228854H1 (BRAMDIT01)

Table 1 (cont.)

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
8	25	2617942	GBLANOT01	1311045F6 (COLNFET02), 1375061F6 (LUNGNOT10), 1580318F6 (DUODNOT01), 2016443F6 (ENDCNOT03), 2617942F6 (GBLANOT01), 2617942H1 (GBLANOT01), 2703625H1 (OVARTUT10)
9	26	2945431	BRAITUT23	782666T6 (MYOMNOT01), 2715384T6 (THYRNOT09), 2945431F6 (BRAITUT23), 2945431H1 (BRAITUT23), 4114919F6 (UTRSTUT07)
10	27	4074113	KIDNNOT26	1754278T6 (LIVRTUT01), 1830341T6 (THP1AZT01), 2122328T6 (BRSTNOT07), 4074113F6 (KIDNNOT26), 4074113H1 (KIDNNOT26)
11	28	1413743	BRAINOT12	1413743H1 (BRAINOT12), 2643096F6 (LUNGTUT08), SBWA04580V1, SBWA04306V1, SBWA00027V1, SBWA00447V1, SBWA00075V1, SBWA03280V1, SBWA00778V1
12	29	1733477	BRSTTUT08	1733477H1 (BRSTTUT08), 1733477F6 (BRSTTUT08), 3346594F6 (BRAITUT24), 1696051F6 (COLNNOT23), SAFC01922F1, 3234341F6 (COLNUCT03), 2634476T6 (COLNTUT15), SBCA03249F1
13	30	2641908	LUNGTUT08	2641908H1 (LUNGTUT08), SASA01215F1, SASB01010F1, SASB01616F1, SASA03523F1
14	31	2656554	LUNGTUT09	2656554H1 (LUNGTUT09), 2656554F6 (LUNGTUT09), 2656554T6 (LUNGTUT09)
15	32	2719228	THYRNOT09	2719228H1 (THYRNOT09), 2719228F6 (THYRNOT09), 532553H1 (BRAINOT03), 1626110F6 (COLNPOT01), 1626488F6 (COLNPOT01), 2100947R6 (BRAITUT02), SBGA05174F1
16	33	3657824	ENDPNOT02	3657824H1 (ENDPNOT02), 2450345F6 (ENDANOT01), 2756551H1 (THP1AZS08), SCBA01425V1, 2394519T6 (THP1AZT01)
17	34	5378485	BRAXNOT01	5378485H1 (BRAXNOT01), SCAA06347V1, 1570276F1 (UTRSNOT05)

Table 2

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Identification/ Homologies	Analytical Methods
1	384	S375 S34 T33 T89 T102 S161 S246 T298		E1-E2 ATPase signatures: R85-D115, T106-I128 P-type cation transporter signature: I109-I128 Transmembrane domains: I130-W351, L271-I288	E1-E2 ATPase [Mus musculus] (GI 2944187)	BLAST BLOCKS MOTIFS PFAM PRINTS
2	846	T80 S106 S22 T40 T109 S110 S141 S229 S245 S255 S288 T408 T614 T679 T707 S819 S24 S274 S293 T386 T450 T536 S568 T581 T610 T673	N301	ABC transporter signatures: F453-L467, G330-Q525, D752-Q816, L744-A758 ATP-binding (P-loop) domains: G337-T344, G659-S666	ABC transporter protein isoform [Homo sapiens] (GI 2522534)	BLAST BLOCKS MOTIFS PFAM PRINTS ProfileScan SPScan
3	511	S183 S99 S385 T387 T39 T106 S240 T376 T461 Y111 Y151 Y194 Y353	N195 N238 N258 N383	AMP-binding enzyme motif: R1-V404 AMP-binding domain signature: K91-M144 Lipocalin signature: E2-M14	Fatty acid transport protein [Homo sapiens] (GI 4206376) Fatty acid transport protein [Mus musculus] (GI 2612939)	BLAST MOTIFS PFAM PRINTS ProfileScan

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Identification/ Homologies	Analytical Methods
4	718	S303 T224 T390 S501 T547 S560 S579 S623 T675 S248 S274 T326 S629 T655	N176 N229 N410	ABC transporter domain: L595-L609 ABC transporter family signatures: G483-V499, L595-D626 ABC transporter transmembrane domain: L128-M408 ABC transporter nucleotide binding domain: G483-G668 ATP/GTP-binding site motif A: G490-T497 Transmembrane domain: F123-V144	ABC transporter protein [Homo sapiens] (GI 4321407) SMDR1 [Schistosoma mansoni] (GI 425474)	BLAST BLOCKS MOTIFS PFAM HMM
5	635	S25 T55 S128 T172 S242 S284 T378 T140 S222 S283 T286 T549	N138 N489 N498 N534	Sodium:solute symporter family signature: M61-A467, N169-A216, T172-V217, G353-I383, I430-A467 Transmembrane domains: M27-Y45, F256-A276, I430-F448, V458-G480	Sodium:solute symporter [Homo sapiens] (GI 4884550)	BLAST BLOCKS PFAM ProfileScan HMM

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Identification/ Homologies	Analytical Methods
6	535	T11 S61 S103 S179 S292 T305 S337 T363 S487 S502 Y109		Amino acid permeases signature: L36-K474 (from PFAM) Transmembrane domains: L76-L96, K111-Y131, S159-N176, I189-V207, N264-Y287, I422-F449, I451-W469	SLC7A8 [Homo sapiens] (GI 4581470) Amino acid transporter (GI 3639058, SEQ ID NO:38)	BLAST BLOCKS MOTIFS PFAM
7	456	T106 S107 S215 S277 T306 T331 S419 S425		Aromatic amino acid permease signature: N53-S98 (from BLOCKS) Transmembrane domains: F47-V69, Y68-L88, T169-L191, L197-I216, S366-Y387, M330-I347, F432-I454	Aromatic amino acid transporter (GI 1840045, SEQ ID NO:39)	BLAST BLOCKS MOTIFS
8	325	T23 T29 S65 T84 S106 S133 S160		Transmembrane domains: I209-D227, I284-T307	UDP-N-acetylglucosamine transporter [Canis familiaris] (GI 3298605)	BLAST MOTIFS
9	178	S4 T22 S132 T151		ABC transporters signature: S2-G141 (from PFAM)	ABC transporter (GI 3335175, SEQ ID NO:41)	BLAST BLOCKS MOTIFS PFAM Profilescan

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Identification/ Homologies	Analytical Methods
10	255	T89 S107 S214		ABC transporters signature: Q45-G227 (from PFAM)	Anthracycline resistance-associated protein (ABC transporter; GI 1279457, SEQ ID NO:42)	BLAST BLOCKS MOTIFS PFAM Profilescan
11	462	T179 S11 S37 T114 S420 T174		Aromatic amino acid permease signatures: F58-A104, A243-F265, L375-K394	Transmembrane transporter [Schizosaccharomyces pombe] (GI 3367790)	BLAST MOTIFS PRINTS
12	758	T174 T251 S376 T615 S616 S633 S678 S713 S733 T138 S574 S603 T656 S687 T726 Y703	N167 N172 N614	Sulfate transporter signature: L198-T508	Transporter [Mus musculus] (GI 5359730) Transporter [Homo sapiens] (GI 291964)	BLAST MOTIFS PFAM
13	336	T32 S99 S254 T274 S103 T104 T133 S273	N116	Transmembrane domains: L211-G230, A175-L196, W42-F61, F139-F159	Organic anion transporter OATP-B [Homo sapiens] (GI 5006263)	BLAST HMM MOTIFS
14	103	S9 S92 Y14	N81	Transmembrane domain: V22-F41	Similar to Human Na+/H+ exchanger 2 (A57644) [Homo sapiens] (GI 1665827)	BLAST HMM MOTIFS

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Identification/ Homologies	Analytical Methods
15	123	S21 S118 T119	N39	Signal peptide: M1-S21	Similar to Sugar transporter [Caenorhabditis elegans] (GI 3878537)	BLAST HMM MOTIFS
16	222	T83 S61 S74			E. coli cation transport protein (GI495778)	BLAST MOTIFS
17	111		N100		Vacuolar proton-ATPase subunit M9.2 [Homo sapiens] (GI 2584789)	BLAST MOTIFS

Table 3

Nucleotide SEQ ID NO:	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
18	Reproductive (0.292) Nervous (0.167) Hematopoietic/Immune (0.125)	Cell proliferative (0.750) Inflammation (0.208)	PSPORT1
19	Reproductive (0.264) Nervous (0.155) Hematopoietic/Immune (0.145)	Cell proliferative (0.691) Inflammation (0.236)	pINCY
20	Reproductive Nervous Gastrointestinal	Cell proliferation (0.66) Inflammation and immune response (0.24)	pINCY
21	Reproductive Nervous Gastrointestinal Hematopoietic/Immune	Cell proliferation (0.73) Inflammation and immune response (0.21)	PSPORT1
22	Nervous Reproductive	Cell proliferation (0.76) Inflammation and immune response (0.33)	pINCY
23	Reproductive (0.333) Nervous (0.296)	Cancer (0.519) Inflammation (0.185)	
24	Reproductive (0.250) Musculoskeletal (0.179) Hematopoietic/Immune (0.143)	Cancer (0.429) Inflammation (0.286)	
25	Reproductive (0.360) Gastrointestinal (0.200) Cardiovascular (0.120)	Cancer (0.520) Inflammation (0.160)	
26	Nervous (0.267) Endocrine (0.133)	Cancer (0.533) Inflammation (0.333)	

Table 3 (cont.)

Nucleotide SEQ ID NO:	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
27	Urologic (0.333) Cardiovascular (0.167) Developmental (0.167)	Cancer (0.667) Inflammation (0.167)	
28	Nervous (0.222) Hematopoietic/Immune (0.185) Reproductive (0.185)	Cell Proliferation (0.481) Inflammation (0.407)	pINCY
29	Gastrointestinal (0.286) Reproductive (0.257) Nervous (0.200)	Cell Proliferation (0.743) Inflammation (0.286)	pINCY
30	Cardiovascular (0.400) Nervous (0.200) Reproductive (0.200) Gastrointestinal (0.200)	Cell Proliferation (0.800) Inflammation (0.200)	pINCY
31	Hematopoietic/Immune (0.500) Reproductive (0.250) Cardiovascular (0.250)	Cell Proliferation (0.750) Inflammation (0.500)	pINCY
32	Gastrointestinal (0.304) Hematopoietic/Immune (0.174) Nervous (0.174)	Cell Proliferation (0.826) Inflammation (0.347)	pINCY
33	Cardiovascular (0.238) Nervous (0.190) Hematopoietic/Immune (0.143)	Cell Proliferation (0.571) Inflammation (0.473)	pINCY
34	Nervous (0.510) Cardiovascular (0.102)	Cell Proliferation (0.612) Inflammation (0.266)	pINCY

Table 4

Nucleotide SEQ ID NO:	Library	Library Comment
18	BRSTTUT03	Library was constructed using RNA isolated from breast tumor tissue removed from a 58-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated multicentric invasive grade 4 lobular carcinoma. The mass was identified in the upper outer quadrant, and three separate nodules were found in the lower outer quadrant of the left breast. Patient history included skin cancer, rheumatic heart disease, osteoarthritis, and tuberculosis. Family history included cerebrovascular disease, coronary artery aneurysm, breast cancer, prostate cancer, atherosclerotic coronary artery disease, and type I diabetes.
19	LUNGTUT12	Library was constructed using RNA isolated from tumorous lung tissue removed from a 70-year-old Caucasian female during a lung lobectomy of the left upper lobe. Pathology indicated grade 3 (of 4) adenocarcinoma and vascular invasion. Patient history included tobacco abuse, depressive disorder, anxiety state, and skin cancer. Family history included cerebrovascular disease, congestive heart failure, colon cancer, depressive disorder, and primary liver cancer.
20	BLADNOT06	Library was constructed using RNA isolated from posterior wall bladder tissue removed from a 66-year-old Caucasian male during a radical prostatectomy, radical cystectomy, and urinary diversion. Pathology for the associated tumor tissue indicated grade 3 transitional cell carcinoma on the anterior wall of the bladder and urothelium. Patient history included lung neoplasm. Family history included a malignant breast neoplasm, tuberculosis, cerebrovascular disease, atherosclerotic coronary artery disease, and lung cancer.
21	PROSNON01	Normalized prostate library was constructed from 4.4 million independent clones from a prostate library. Starting RNA was made from prostate tissue removed from a 28-year-old Caucasian male. The normalization and hybridization conditions were adapted from Soares, M.B. et al. (1994) Proc. Natl. Acad. Sci. USA 91:9228-9232, using a longer (19 hour) reannealing hybridization period.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Comment
22	BRSTTUT14	Library was constructed using RNA isolated from breast tumor tissue removed from a 62-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated an invasive grade 3 (of 4), nuclear grade 3 (of 3) adenocarcinoma, ductal type. Ductal carcinoma in situ, comedo type, comprised 60% of the tumor mass. Metastatic adenocarcinoma was identified in one (of 14) axillary lymph nodes with no perinodal extension. Tumor cells were strongly positive for estrogen receptors and weakly positive for progesterone receptors. Patient history included benign colon neoplasm, hyperlipidemia, cardiac dysrhythmia, and obesity. Family history included atherosclerotic coronary artery disease, myocardial infarction, colon cancer, ovarian cancer, lung cancer, and cerebrovascular disease.
23	HNT3AZT01	Library was constructed using RNA isolated from the hNT2 cell line (derived from a human teratocarcinoma that exhibited properties characteristic of a committed neuronal precursor). Cells were treated for three days with 0.35 micromolar 5-aza-2'-deoxycytidine (AZ).
24	THYRNOT08	Library was constructed using RNA isolated from the diseased left thyroid tissue removed from a 13-year-old Caucasian female during a complete thyroidectomy. Pathology indicated lymphocytic thyroiditis.
25	GBLANOT01	Library was constructed using RNA isolated from diseased gallbladder tissue removed from a 53-year-old Caucasian female during a cholecystectomy. Pathology indicated mild chronic cholecystitis and cholelithiasis with approximately 150 mixed gallstones. Family history included benign hypertension.
26	BRAITUT23	Library was constructed using RNA isolated from left posterior brain tumor tissue removed from a 36-year-old male during a cerebral meninges lesion excision. Pathology indicated a meningioma. Family history included malignant skin melanoma, atherosclerotic coronary artery disease, hyperlipidemia, Huntington's chorea, and rheumatoid arthritis.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Comment
27	KIDNNOT26	Library was constructed using RNA isolated from left kidney medulla and cortex tissue removed from a 53-year-old Caucasian female during a nephroureterectomy. Patient history included hyperlipidemia, cardiac dysrhythmia, metrorrhagia, cerebrovascular disease, and atherosclerotic coronary artery disease.
28	BRAINOT12	Library was constructed using RNA isolated from brain tissue removed from the right frontal lobe of a 5-year-old Caucasian male during a hemispherectomy. Pathology indicated extensive polymicrogyria and mild to moderate gliosis (predominantly subpial and subcortical), which are consistent with chronic seizure disorder. The patient presented with intractable convulsive epilepsy. Family history included a cervical neoplasm.
29	BRSTTUT08	Library was constructed using RNA isolated from breast tumor tissue removed from a 45-year-old Caucasian female during unilateral extended simple mastectomy. Pathology indicated invasive nuclear grade 2-3 adenocarcinoma (ductal type), with 3 of 23 lymph nodes positive for metastatic disease. Greater than 50% of the tumor volume was in-situ, both comedo and non-comedo types. There were also positive estrogen/progesterone receptors and uninvolved tissue showing proliferative changes. Patient history included valvuloplasty of mitral valve without replacement, rheumatic mitral insufficiency, and rheumatic heart disease. Family history included acute myocardial infarction and atherosclerotic coronary artery disease and type II diabetes.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Comment
30	LUNGTUT08	Library was constructed using RNA isolated from lung tumor tissue removed from a 63-year-old Caucasian male during a right upper lobectomy with fiberoptic bronchoscopy. Pathology indicated a grade 3 adenocarcinoma,forming a mass penetrating the pleura of the right upper lung lobe. The bronchial margin of the resection was free of tumor. The tissue from the superior segment of the right lower lobe lung showed multiple (2) calcified granulomas. Multiple lymph nodes were negative for tumor. In addition, budding yeast forms resembling histoplasma identified in two lymph nodes stained with GMS (silver). Patient history included atherosclerotic coronary artery disease, an acute myocardial infarction, rectal cancer, an asymtomatic abdominal aortic aneurysm, tobacco abuse, and cardiac dysrhythmia. Family history included congestive heart failure, stomach cancer, lung cancer, type II diabetes, atherosclerotic coronary artery disease, and acute myocardial infarction.
31	LUNGTUT09	Library was constructed RNA isolated from lung tumor tissue removed from a 68-year-old Caucasian male during segmental lung resection. Pathology indicated invasive grade 3 squamous cell carcinoma in the right upper lobe, forming an infiltrating mass involving the bronchus and the surrounding parenchyma. One (of 4) intrapulmonary peribronchial lymph nodes contained a metastatic tumor. An apical cap was identified. One (of 15) right paratracheal lower lymph nodes contained a metastatic tumor. Permanent superior mediastinal sections revealed metastatic squamous cell carcinoma in the lymph nodes. Patient history included of type II diabetes, thyroid disorder, depressive disorder, hyperlipidemia, esophageal ulcer, and tobacco use. Family history included brain cancer and atherosclerotic coronary artery disease.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Comment
32	THYRNOT09	Library was constructed using RNA isolated from diseased thyroid tissue removed from an 18-year-old Caucasian female during an unilateral thyroid lobectomy and regional lymph node excision. Pathology indicated adenomatous goiter. This was associated with a follicular adenoma of the thyroid. The right neck lymph nodes displayed reactive follicular hyperplasia. The patient presented with hypophosphatemia. Patient history included compression of the brain, headache, iron deficiency anemia, active rickets, epidermal nevus syndrome (treated with lasers), and osteitis deformans. Family history included thyroid cancer, type II diabetes, benign hypertension, and nodular lymphoma.
33	ENDPNOT02	Library was constructed using RNA isolated from pulmonary artery endothelial cells removed from a 10-year-old Caucasian male. The cells were treated with TNF alpha and IL-1 beta 10ng/ml each for 20 hours.
34	BRAXNOT01	Library was constructed using RNA isolated from cerebellar tissue removed from a 70-year-old male. Patient history included chronic obstructive airways disease and left ventricular failure.

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25: 3389-3402.	*ESTs:* Probability value= 1.0E-8 or less *Full Length sequences:* Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	*ESTs:* fasta E value=1.06E-6 *Assembled ESTs:* fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less *Full Length sequences:* fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and, if applicable, Probability value= 1.0E-3 or less[1]
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score=10-50 bits for PFAM hits, depending on individual protein families

WO 00/26345 PCT/US99/26048

Table 5 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Normalized quality score≥GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score=3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. supra; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. A substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-17 and fragments thereof.

2. A substantially purified variant having at least 90% amino acid sequence identity to the amino acid sequence of claim 1.

3. An isolated and purified polynucleotide encoding the polypeptide of claim 1.

4. An isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide of claim 3.

5. An isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide of claim 3.

6. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 3.

7. A method for detecting a polynucleotide, the method comprising the steps of:

(a) hybridizing the polynucleotide of claim 6 to at least one nucleic acid in a sample, thereby forming a hybridization complex; and

(b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of the polynucleotide in the sample.

8. The method of claim 7 further comprising amplifying the polynucleotide prior to hybridization.

9. An isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:18-34 and fragments thereof.

10. An isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide of claim 9.

11. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 9.

12. An expression vector comprising at least a fragment of the polynucleotide of claim 3.

13. A host cell comprising the expression vector of claim 12.

14. A method for producing a polypeptide, the method comprising the steps of:

a) culturing the host cell of claim 13 under conditions suitable for the expression of the polypeptide; and

b) recovering the polypeptide from the host cell culture.

15. A pharmaceutical composition comprising the polypeptide of claim 1 in conjunction with a suitable pharmaceutical carrier.

16. A purified antibody which specifically binds to the polypeptide of claim 1.

17. A purified agonist of the polypeptide of claim 1.

18. A purified antagonist of the polypeptide of claim 1.

19. A method for treating or preventing a disorder associated with decreased expression or activity of MTRP, the method comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 15.

20. A method for treating or preventing a disorder associated with increased expression or activity of MTRP, the method comprising administering to a subject in need of such treatment an effective amount of the antagonist of claim 18.

```
1    M-----------------------------  1720440
1    MRAPGAGTASVASLALLWFLGLPWTWSAAA  GI 2612939

2    ------------------------------  1720440
31   AFCVYVGGGGWRFLRIVCKTARRDLFGLSV  GI 2612939

2    ------------------------------  1720440
61   LIRVRLELRRHRRAGDTIPCIFQAVARRQP  GI 2612939

2    ------------------------------  1720440
91   ERLALVDASSGICWTFAQLDTYSNAVANLF  GI 2612939

2    ---------------ENRNEFVGLWLGMAK  1720440
121  RQLGFAPGDVVAVFLEGRPEFVGLWLGLAK  GI 2612939

17   LGVEAALINTNLRRDALLHCLTTSRARALV  1720440
151  AGVVAALLNVNLRREPLAFCLGTSAAKALI  GI 2612939
```

FIGURE 1A

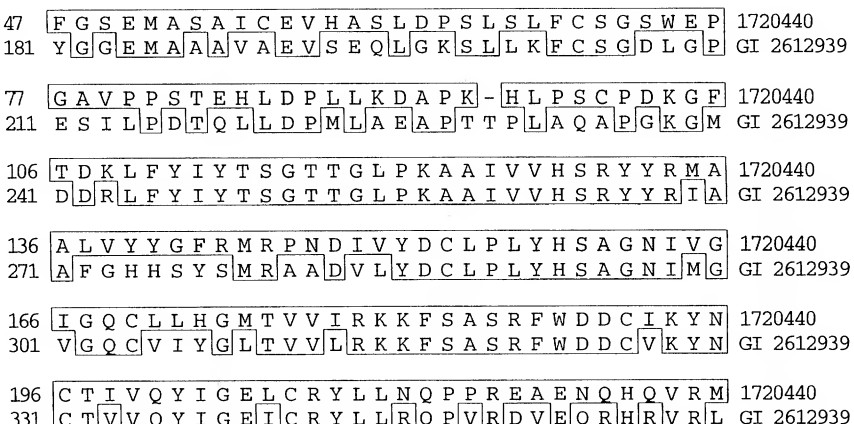
47 F G S E M A S A I C E V H A S L D P S L S L F C S G S W E P 1720440
181 Y G G E M A A A V A E V S E Q L G K S L L K F C S G D L G P GI 2612939
77 G A V P P S T E H L D P L L K D A P K - H L P S C P D K G F 1720440
211 E S I L P D T Q L L D P M L A E A P T T P L A Q A P G K G M GI 2612939
106 T D K L F Y I Y T S G T T G L P K A A I V V H S R Y Y R M A 1720440
241 D D R L F Y I Y T S G T T G L P K A A I V V H S R Y Y R I A GI 2612939
136 A L V Y Y G F R M R P N D I V Y D C L P L Y H S A G N I V G 1720440
271 A F G H H S Y S M R A A D V L Y D C L P L Y H S A G N I M G GI 2612939
166 I G Q C L L H G M T V V I R K K F S A S R F W D D C I K Y N 1720440
301 V G Q C V I Y G L T V V L R K K F S A S R F W D D C V K Y N GI 2612939
196 C T I V Q Y I G E L C R Y L L N Q P P R E A E N Q H Q V R M 1720440
331 C T V V Q Y I G E I C R Y L L R Q P V R D V E Q R H R V R L GI 2612939

FIGURE 1B

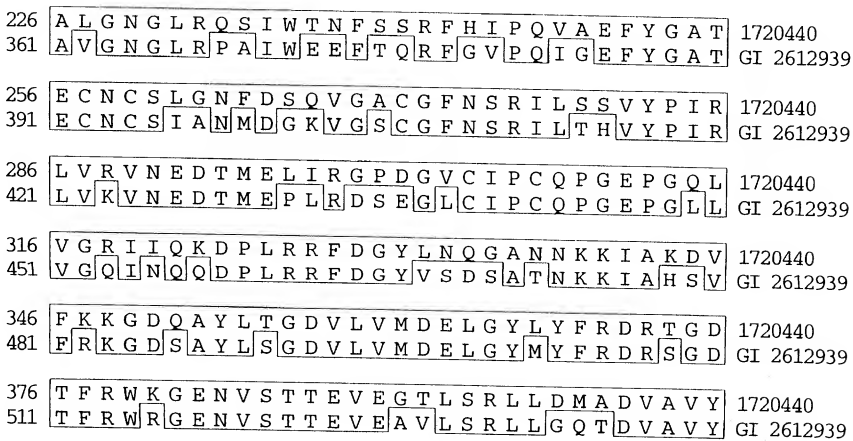
226 A L G N G L R Q S I W T N F S S R F H I P Q V A E F Y G A T 1720440
361 A V G N G L R P A I W E E F T Q R F G V P Q I G E F Y G A T GI 2612939
256 E C N C S L G N F D S Q V G A C G F N S R I L S S V Y P I R 1720440
391 E C N C S I A N M D G K V G S C G F N S R I L T H V Y P I R GI 2612939
286 L V R V N E D T M E L I R G P D G V C I P C Q P G E P G Q L 1720440
421 L V K V N E D T M E P L R D S E G L C I P C Q P G E P G L L GI 2612939
316 V G R I I Q K D P L R R F D G Y L N Q G A N N K K I A K D V 1720440
451 V G Q I N Q Q D P L R R F D G Y V S D S A T N K K I A H S V GI 2612939
346 F K K G D Q A Y L T G D V L V M D E L G Y L Y F R D R T G D 1720440
481 F R K G D S A Y L S G D V L V M D E L G Y M Y F R D R S G D GI 2612939
376 T F R W K G E N V S T T E V E G T L S R L L D M A D V A V Y 1720440
511 T F R W R G E N V S T T E V E A V L S R L L G Q T D V A V Y GI 2612939

FIGURE 1C

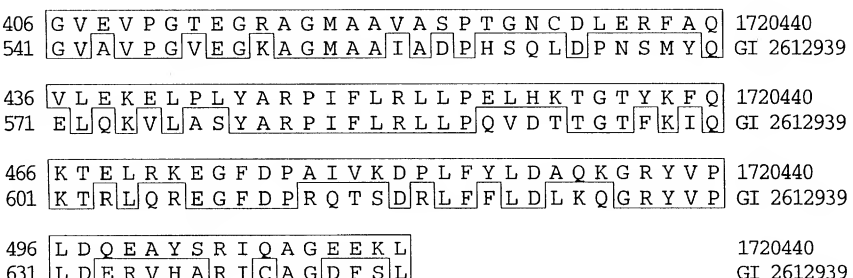
406 G V E V P G T E G R A G M A A V A S P T G N C D L E R F A Q 1720440
541 G V A V P G V E G K A G M A A I A D P H S Q L D P N S M Y Q GI 2612939
436 V L E K E L P L Y A R P I F L R L L P E L H K T G T Y K F Q 1720440
571 E L Q K V L A S Y A R P I F L R L L P Q V D T T G T F K I Q GI 2612939
466 K T E L R K E G F D P A I V K D P L F Y L D A Q K G R Y V P 1720440
601 K T R L Q R E G F D P R Q T S D R L F F L D L K Q G R Y V P GI 2612939
496 L D Q E A Y S R I Q A G E E K L 1720440
631 L D E R V H A R I C A G D F S L GI 2612939

FIGURE 1D

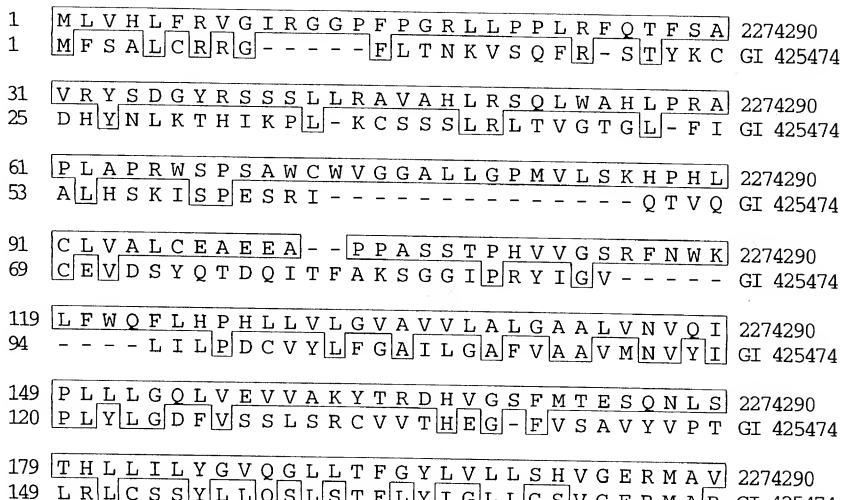
1 MLVHLFRVGIRGGPFPGRLLPPLRFQTFSA 2274290
1 MFSALCRRG-----FLTNKVSQFR-STYKC GI 425474
31 VRYSDGYRSSSLLRAVAHLRSQLWAHLPRA 2274290
25 DHYNLKTHIKPL-KCSSSLRLTVGTGL-FI GI 425474
61 PLAPRWSPSAWCWVGGALLGPMVLSKHPHL 2274290
53 ALHSKISPESRI----------QTVQ GI 425474
91 CLVALCEAEEA--PPASSTPHVVGSRFNWK 2274290
69 CEVDSYQTDQITFAKSGGIPRYIGV----- GI 425474
119 LFWQFLHPHLLVLGVAVVLALGAALVNVQI 2274290
94 ----LILPDCVYLFGAILGAFVAAVMNVYI GI 425474
149 PLLLGQLVEVVAKYTRDHVGSFMTESQNLS 2274290
120 PLYLGDFVSSLSRCVVTHEG-FVSAVYVPT GI 425474
179 THLLILYGVQGLLTFGYLVLLSHVGERMAV 2274290
149 LRLCSSYLLQSLSTFLYIGLLGSVGERMAR GI 425474

FIGURE 2A

```
209 DMRRALFSSLLRQDITFFDANKTGQLVSRL 2274290
179 RMRIQLFRKLVYQDVAYFDVHSSGKLVEII GI 425474

239 TTDVQEFKSSFKLVISQGLRSCTQVAGCLV 2274290
209 GSDVQNFKSSFKQCISQGLRNGIQVVGSVF GI 425474

269 SLSMLSTRLTLLLMVATPALMGVGTLMGSG 2274290
239 ALLSISPTLTAALIGCLPCVFLIGSLMGTE GI 425474

299 LRKLSRQCQEQIARAMGVADEALGNVRTVR 2274290
269 LRHISREVQSQNSLFASLIDEAFSHIRTVK GI 425474

329 AFAMEQREEERYGAELEACRCRAEELGRGI 2274290
299 SLAMEDFLINKINYNVDKAKMLSEKLSFGI GI 425474

359 ALFQGLSNIAFNCMVLGTLFIGGSLVAGQQ 2274290
329 GSFQGLSNLTLNGVVLGVLYVGGHLMSRGE GI 425474
```

FIGURE 2B

```
389 LTGGDLMSFLVASQTVQRSMANLSVLFGQV 2274290
359 LDAGHLMSFLATTQTLQRSLTQLSLLYGQV GI 425474

419 VRGLSAGARVFEYMALNP---CIPLSGGC- 2274290
389 VRGYTALKRIHDILALPSGIGSIPSSSSSL GI 425474

445 CVPKEQLRG-------------SVTFQNVC 2274290
419 VVSKQHVNNIKELPSSSIYSAPSIEFSDVK GI 425474

462 FSYPCRPGFEVLKDFTLTLPPGKIVALVGQ 2274290
449 FAYPNRPETIVLNELSMFLPGGKVIALVGQ GI 425474

492 SGGGKTTVASLLERFYDPTAGVVMLDGRDL 2274290
479 SGAGKSTVVSLLERFYDPISGEILLNGDKL GI 425474

522 RTLDPSWLRGQVVGFISQEPVLFGTTIMEN 2274290
509 TNFNVNYLRSKLIGYISQEPQIFNASIRENI GI 425474
```

FIGURE 2C

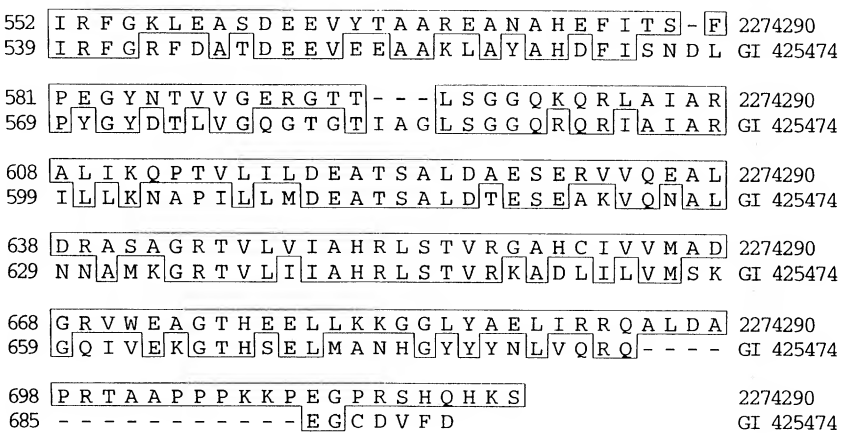
552 I R F G K L E A S D E E V Y T A A R E A N A H E F I T S - F 2274290
539 I R F G R F D A T D E E V E E A A K L A Y A H D F I S N D L GI 425474
581 P E G Y N T V V G E R G T T - - - L S G G Q K Q R L A I A R 2274290
569 P Y G Y D T L V G Q G T G T I A G L S G G Q R Q R I A I A R GI 425474
608 A L I K Q P T V L I L D E A T S A L D A E S E R V V Q E A L 2274290
599 I L L K N A P I L L M D E A T S A L D T E S E A K V Q N A L GI 425474
638 D R A S A G R T V L V I A H R L S T V R G A H C I V V M A D 2274290
629 N N A M K G R T V L I I A H R L S T V R K A D L I L V M S K GI 425474
668 G R V W E A G T H E E L L K K G G L Y A E L I R R Q A L D A 2274290
659 G Q I V E K G T H S E L M A N H G Y Y Y N L V Q R Q - - - - GI 425474
698 P R T A A P P P K K P E G P R S H Q H K S 2274290
685 - - - - - - - - - - - E G C D V F D GI 425474

FIGURE 2D

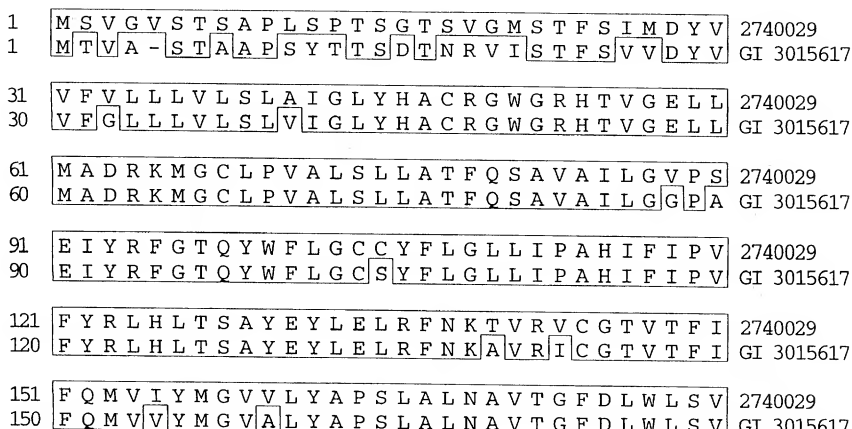
1 MSVGVSTSAPLSPTSGTSVGMSTFSIMDYV 2740029
1 MTVA-STAAPSYTTSDTNRVISTFSVVDYV GI 3015617
31 VFVLLLVLSLAIGLYHACRGWGRHTVGELL 2740029
30 VFGLLLVLSLVIGLYHACRGWGRHTVGELL GI 3015617
61 MADRKMGCLPVALSLLATFQSAVAILGVPS 2740029
60 MADRKMGCLPVALSLLATFQSAVAILGGPA GI 3015617
91 EIYRFGTQYWFLGCCYFLGLLIPAHIFIPV 2740029
90 EIYRFGTQYWFLGCSYFLGLLIPAHIFIPV GI 3015617
121 FYRLHLTSAYEYLELRFNKTVRVCGTVTFI 2740029
120 FYRLHLTSAYEYLELRFNKAVRICGTVTFI GI 3015617
151 FQMVIYMGVVLYAPSLALNAVTGFDLWLSV 2740029
150 FQMVVYMGVALYAPSLALNAVTGFDLWLSV GI 3015617

FIGURE 3A

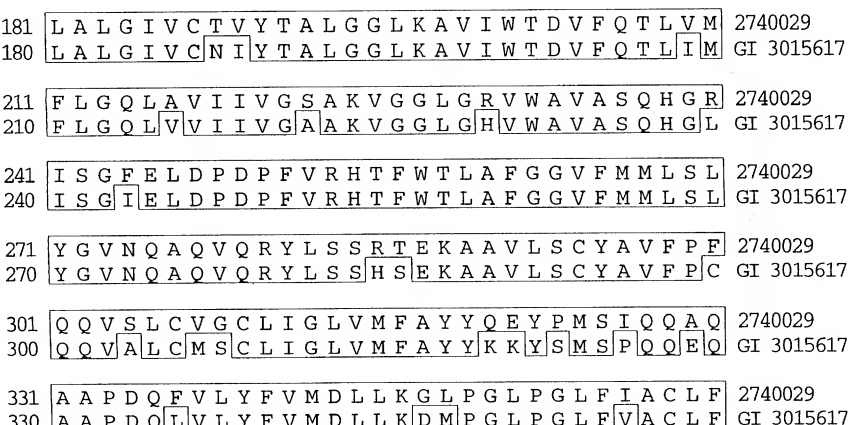
181 L A L G I V C T V Y T A L G G L K A V I W T D V F Q T L V M 2740029
180 L A L G I V C N I Y T A L G G L K A V I W T D V F Q T L I M GI 3015617
211 F L G Q L A V I I V G S A K V G G L G R V W A V A S Q H G R 2740029
210 F L G Q L V V I I V G A A K V G G L G H V W A V A S Q H G L GI 3015617
241 I S G F E L D P D P F V R H T F W T L A F G G V F M M L S L 2740029
240 I S G I E L D P D P F V R H T F W T L A F G G V F M M L S L GI 3015617
271 Y G V N Q A Q V Q R Y L S S R T E K A A V L S C Y A V F P F 2740029
270 Y G V N Q A Q V Q R Y L S S H S E K A A V L S C Y A V F P C GI 3015617
301 Q Q V S L C V G C L I G L V M F A Y Y Q E Y P M S I Q Q A Q 2740029
300 Q Q V A L C M S C L I G L V M F A Y Y K K Y S M S P Q Q E Q GI 3015617
331 A A P D Q F V L Y F V M D L L K G L P G L P G L F I A C L F 2740029
330 A A P D Q L V L Y F V M D L L K D M P G L P G L F V A C L F GI 3015617

FIGURE 3B

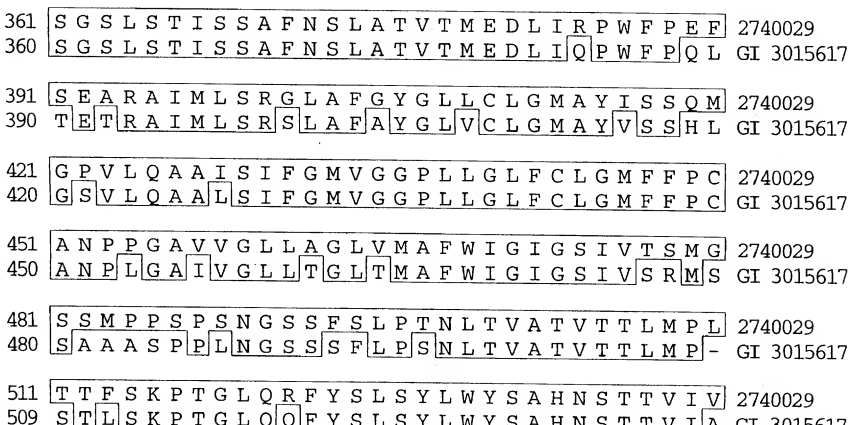
361 S G S L S T I S S A F N S L A T V T M E D L I R P W F P E F 2740029
360 S G S L S T I S S A F N S L A T V T M E D L I Q P W F P Q L GI 3015617
391 S E A R A I M L S R G L A F G Y G L L C L G M A Y I S S Q M 2740029
390 T E T R A I M L S R S L A F A Y G L V C L G M A Y V S S H L GI 3015617
421 G P V L Q A A I S I F G M V G G P L L G L F C L G M F F P C 2740029
420 G S V L Q A A L S I F G M V G G P L L G L F C L G M F F P C GI 3015617
451 A N P P G A V V G L L A G L V M A F W I G I G S I V T S M G 2740029
450 A N P L G A I V G L L T G L T M A F W I G I G S I V S R M S GI 3015617
481 S S M P P S P S N G S S F S L P T N L T V A T V T T L M P L 2740029
480 S A A A S P P L N G S S S F L P S N L T V A T V T T L M P - GI 3015617
511 T T F S K P T G L Q R F Y S L S Y L W Y S A H N S T T V I V 2740029
509 S T L S K P T G L Q Q F Y S L S Y L W Y S A H N S T T V I A GI 3015617

FIGURE 3C

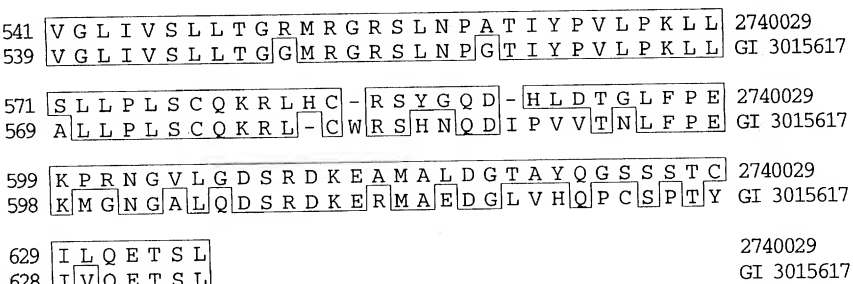
541 VGLIVSLLTGRMRGRSLNPATIYPVLPKLL 2740029
539 VGLIVSLLTGGMRGRSLNPGTIYPVLPKLL GI 3015617
571 SLLPLSCQKRLHC-RSYGQD-HLDTGLFPE 2740029
569 ALLPLSCQKRL-CWRSHNQDIPVVTNLFPE GI 3015617
599 KPRNGVLGDSRDKEAMALDGTAYQGSSSTC 2740029
598 KMGNGALQDSRDKERMAEDGLVHQPCSPTY GI 3015617
629 ILQETSL 2740029
628 IVQETSL GI 3015617

FIGURE 3D

SEQUENCE LISTING

<110> INCYTE PHARMACEUTICALS, INC.
HILLMAN, Jennifer L.
YUE, Henry
TANG, Y. Tom
LAL, Preeti
CORLEY, Neil C.
GUEGLER, Karl J.
BAUGHN, Mariah R.
AZIMZAI, Yalda
LU, Dyung Aina M.

<120> MEMBRANE TRANSPORT PROTEINS

<130> PF-0633 PCT

<140> To Be Assigned
<141> Herewith

<150> 09/186,778; unassigned; 09/200,277; unassigned; 09/221,405; unassigned; 60/121,896
<151> 1998-11-04; 1998-11-04; 1998-11-24; 1998-11-24; 1998-12-22; 1998-12-22; 1999-02-26

<160> 43

<170> PERL Program

<210> 1
<211> 384
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 961344CD1

<400> 1

```
Met Leu Thr Gly Asp Lys Leu Glu Thr Ala Thr Cys Ile Ala Lys
  1               5                  10                  15
Ser Ser His Leu Val Ser Arg Thr Gln Asp Ile His Ile Phe Arg
             20                  25                  30
Gln Val Thr Ser Arg Gly Glu Ala His Leu Glu Leu Asn Ala Phe
             35                  40                  45
Arg Arg Lys His Asp Cys Ala Leu Val Ile Ser Gly Asp Ser Leu
             50                  55                  60
Glu Val Cys Leu Lys Tyr Tyr Glu His Glu Phe Val Glu Leu Ala
             65                  70                  75
Cys Gln Cys Pro Ala Val Val Cys Cys Arg Cys Ser Pro Thr Gln
             80                  85                  90
Lys Ala Arg Ile Val Thr Leu Leu Gln Gln His Thr Gly Arg Arg
             95                 100                 105
```

```
Thr Cys Ala Ile Gly Asp Gly Gly Asn Asp Val Ser Met Ile Gln
                110                 115                 120
Ala Ala Asp Cys Gly Ile Gly Ile Glu Gly Lys Glu Gly Lys Gln
                125                 130                 135
Ala Ser Leu Ala Ala Asp Phe Ser Ile Thr Gln Phe Arg His Ile
                140                 145                 150
Gly Arg Leu Leu Met Val His Gly Arg Asn Ser Tyr Lys Arg Ser
                155                 160                 165
Ala Ala Leu Gly Gln Phe Val Met His Arg Gly Leu Ile Ile Ser
                170                 175                 180
Thr Met Gln Ala Val Phe Ser Ser Val Phe Tyr Phe Ala Ser Val
                185                 190                 195
Pro Leu Tyr Gln Gly Phe Leu Met Val Gly Tyr Ala Thr Ile Tyr
                200                 205                 210
Thr Met Phe Pro Val Phe Ser Leu Val Leu Asp Gln Asp Val Lys
                215                 220                 225
Pro Glu Met Ala Met Leu Tyr Pro Glu Leu Tyr Lys Asp Leu Thr
                230                 235                 240
Lys Gly Arg Ser Leu Ser Phe Lys Thr Phe Leu Ile Trp Val Leu
                245                 250                 255
Ile Ser Ile Tyr Gln Gly Gly Ile Leu Met Tyr Gly Ala Leu Val
                260                 265                 270
Leu Phe Glu Ser Glu Phe Val His Val Val Ala Ile Ser Phe Thr
                275                 280                 285
Ala Leu Ile Leu Thr Glu Leu Leu Met Val Ala Leu Thr Val Arg
                290                 295                 300
Thr Trp His Trp Leu Met Val Val Ala Glu Phe Leu Ser Leu Gly
                305                 310                 315
Cys Tyr Val Ser Ser Leu Ala Phe Leu Asn Glu Tyr Phe Gly Ile
                320                 325                 330
Gly Arg Val Ser Phe Gly Ala Phe Leu Asp Val Ala Phe Ile Thr
                335                 340                 345
Thr Val Thr Phe Leu Trp Lys Val Ser Ala Ile Thr Val Val Ser
                350                 355                 360
Cys Leu Pro Leu Tyr Val Leu Lys Tyr Leu Arg Arg Lys Leu Ser
                365                 370                 375
Pro Pro Ser Tyr Cys Lys Leu Ala Ser
                380
```

<210> 2
<211> 846
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3128782CD1

<400> 2

```
Met Pro Lys Ala Pro Lys Gln Gln Pro Pro Glu Pro Glu Trp Ile
  1               5                  10                  15
Gly Asp Gly Glu Ser Thr Ser Pro Ser Asp Lys Val Val Lys Lys
                20                  25                  30
Gly Lys Lys Asp Lys Lys Ile Lys Lys Thr Phe Phe Glu Glu Leu
```

```
                35                  40                  45
Ala Val Glu Asp Lys Gln Ala Gly Glu Glu Glu Lys Val Leu Lys
            50                  55                  60
Glu Lys Glu Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Lys
        65                  70                  75
Lys Lys Arg Asp Thr Arg Lys Gly Arg Arg Lys Lys Asp Val Asp
                80                  85                  90
Asp Asp Gly Glu Glu Lys Glu Leu Met Glu Arg Leu Lys Lys Leu
            95                  100                 105
Ser Val Pro Thr Ser Asp Glu Glu Asp Glu Val Pro Ala Pro Lys
        110                 115                 120
Pro Arg Gly Gly Lys Lys Thr Lys Gly Gly Asn Val Phe Ala Ala
    125                 130                 135
Leu Ile Gln Asp Gln Ser Glu Glu Glu Glu Glu Glu Glu Lys His
140                 145                 150
Pro Pro Lys Pro Ala Lys Pro Glu Lys Asn Arg Ile Asn Lys Ala
155                 160                 165
Val Ser Glu Glu Gln Gln Pro Ala Leu Lys Gly Lys Lys Gly Lys
            170                 175                 180
Glu Glu Lys Ser Lys Gly Lys Ala Lys Pro Gln Asn Lys Phe Ala
        185                 190                 195
Ala Leu Asp Asn Glu Glu Glu Asp Lys Glu Glu Glu Ile Ile Lys
    200                 205                 210
Glu Lys Glu Pro Pro Lys Gln Gly Lys Glu Lys Ala Lys Lys Ala
215                 220                 225
Glu Gln Gly Ser Glu Glu Glu Gly Glu Gly Glu Glu Glu Glu Glu
            230                 235                 240
Glu Gly Gly Glu Ser Lys Ala Asp Asp Pro Tyr Ala His Leu Ser
        245                 250                 255
Lys Lys Glu Lys Lys Leu Lys Lys Gln Met Glu Tyr Glu Arg
    260                 265                 270
Gln Val Ala Ser Leu Lys Ala Ala Asn Ala Ala Glu Asn Asp Phe
275                 280                 285
Ser Val Ser Gln Ala Glu Met Ser Ser Arg Gln Ala Met Leu Glu
            290                 295                 300
Asn Ala Ser Asp Ile Lys Leu Glu Lys Phe Ser Ile Ser Ala His
        305                 310                 315
Gly Lys Glu Leu Phe Val Asn Ala Asp Leu Tyr Ile Val Ala Gly
    320                 325                 330
Arg Arg Tyr Gly Leu Val Gly Pro Asn Gly Lys Gly Lys Thr Thr
335                 340                 345
Leu Leu Lys His Ile Ala Asn Arg Ala Leu Ser Ile Pro Pro Asn
            350                 355                 360
Ile Asp Val Leu Leu Cys Glu Gln Glu Val Val Ala Asp Glu Thr
        365                 370                 375
Pro Ala Val Gln Ala Val Leu Arg Ala Asp Thr Lys Arg Leu Lys
    380                 385                 390
Leu Leu Glu Glu Glu Arg Arg Leu Gln Gly Gln Leu Glu Gln Gly
395                 400                 405
Asp Asp Thr Ala Ala Glu Arg Leu Glu Lys Val Tyr Glu Glu Leu
            410                 415                 420
Arg Ala Thr Gly Ala Ala Ala Ala Glu Ala Lys Ala Arg Arg Ile
        425                 430                 435
Leu Ala Gly Leu Gly Phe Asp Pro Glu Met Gln Asn Arg Pro Thr
    440                 445                 450
```

Gln Lys Phe Ser Gly Gly Trp Arg Met Arg Val Ser Leu Ala Arg
455 460 465
Ala Leu Phe Met Glu Pro Thr Leu Leu Met Leu Asp Glu Pro Thr
470 475 480
Asn His Leu Asp Leu Asn Ala Val Ile Trp Leu Asn Asn Tyr Leu
485 490 495
Gln Gly Trp Arg Lys Thr Leu Leu Ile Val Ser His Asp Gln Gly
500 505 510
Phe Leu Asp Asp Val Cys Thr Asp Ile Ile His Leu Asp Ala Gln
515 520 525
Arg Leu His Tyr Tyr Arg Gly Asn Tyr Met Thr Phe Lys Lys Met
530 535 540
Tyr Gln Gln Lys Gln Lys Glu Leu Leu Lys Gln Tyr Glu Lys Gln
545 550 555
Glu Lys Lys Leu Lys Glu Leu Lys Ala Gly Gly Lys Ser Thr Lys
560 565 570
Gln Ala Glu Lys Gln Thr Lys Glu Ala Leu Thr Arg Lys Gln Gln
575 580 585
Lys Cys Arg Arg Lys Asn Gln Asp Glu Glu Ser Gln Glu Ala Pro
590 595 600
Glu Leu Leu Lys Arg Pro Lys Glu Tyr Thr Val Arg Phe Thr Phe
605 610 615
Pro Asp Pro Pro Pro Leu Ser Pro Pro Val Leu Gly Leu His Gly
620 625 630
Val Thr Phe Gly Tyr Gln Gly Gln Lys Pro Leu Phe Lys Asn Leu
635 640 645
Asp Phe Gly Ile Asp Met Asp Ser Arg Ile Cys Ile Val Gly Pro
650 655 660
Asn Gly Val Gly Lys Ser Thr Leu Leu Leu Leu Leu Thr Gly Lys
665 670 675
Leu Thr Pro Thr His Gly Glu Met Arg Lys Asn His Arg Leu Lys
680 685 690
Ile Gly Phe Phe Asn Gln Gln Tyr Ala Glu Gln Leu Arg Met Glu
695 700 705
Glu Thr Pro Thr Glu Tyr Leu Gln Arg Gly Phe Asn Leu Pro Tyr
710 715 720
Gln Asp Ala Arg Lys Cys Leu Gly Arg Phe Gly Leu Glu Ser His
725 730 735
Ala His Thr Ile Gln Ile Cys Lys Leu Ser Gly Gly Gln Lys Ala
740 745 750
Arg Val Val Phe Ala Glu Leu Ala Cys Arg Glu Pro Asp Val Leu
755 760 765
Ile Leu Asp Glu Pro Thr Asn Asn Leu Asp Ile Glu Ser Ile Asp
770 775 780
Ala Leu Gly Glu Ala Ile Asn Glu Tyr Lys Gly Ala Val Ile Val
785 790 795
Val Ser His Asp Ala Arg Leu Ile Thr Glu Thr Asn Cys Gln Leu
800 805 810
Trp Val Val Glu Glu Gln Ser Val Ser Gln Ile Asp Gly Asp Phe
815 820 825
Glu Asp Tyr Lys Arg Glu Val Leu Glu Ala Leu Gly Glu Val Met
830 835 840
Val Ser Arg Pro Arg Glu
845

<210> 3
<211> 511
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1720440CD1

<400> 3
Met Glu Asn Arg Asn Glu Phe Val Gly Leu Trp Leu Gly Met Ala
1 5 10 15
Lys Leu Gly Val Glu Ala Ala Leu Ile Asn Thr Asn Leu Arg Arg
20 25 30
Asp Ala Leu Leu His Cys Leu Thr Thr Ser Arg Ala Arg Ala Leu
35 40 45
Val Phe Gly Ser Glu Met Ala Ser Ala Ile Cys Glu Val His Ala
50 55 60
Ser Leu Asp Pro Ser Leu Ser Leu Phe Cys Ser Gly Ser Trp Glu
65 70 75
Pro Gly Ala Val Pro Pro Ser Thr Glu His Leu Asp Pro Leu Leu
80 85 90
Lys Asp Ala Pro Lys His Leu Pro Ser Cys Pro Asp Lys Gly Phe
95 100 105
Thr Asp Lys Leu Phe Tyr Ile Tyr Thr Ser Gly Thr Thr Gly Leu
110 115 120
Pro Lys Ala Ala Ile Val Val His Ser Arg Tyr Tyr Arg Met Ala
125 130 135
Ala Leu Val Tyr Tyr Gly Phe Arg Met Arg Pro Asn Asp Ile Val
140 145 150
Tyr Asp Cys Leu Pro Leu Tyr His Ser Ala Gly Asn Ile Val Gly
155 160 165
Ile Gly Gln Cys Leu Leu His Gly Met Thr Val Val Ile Arg Lys
170 175 180
Lys Phe Ser Ala Ser Arg Phe Trp Asp Asp Cys Ile Lys Tyr Asn
185 190 195
Cys Thr Ile Val Gln Tyr Ile Gly Glu Leu Cys Arg Tyr Leu Leu
200 205 210
Asn Gln Pro Pro Arg Glu Ala Glu Asn Gln His Gln Val Arg Met
215 220 225
Ala Leu Gly Asn Gly Leu Arg Gln Ser Ile Trp Thr Asn Phe Ser
230 235 240
Ser Arg Phe His Ile Pro Gln Val Ala Glu Phe Tyr Gly Ala Thr
245 250 255
Glu Cys Asn Cys Ser Leu Gly Asn Phe Asp Ser Gln Val Gly Ala
260 265 270
Cys Gly Phe Asn Ser Arg Ile Leu Ser Ser Val Tyr Pro Ile Arg
275 280 285
Leu Val Arg Val Asn Glu Asp Thr Met Glu Leu Ile Arg Gly Pro
290 295 300
Asp Gly Val Cys Ile Pro Cys Gln Pro Gly Glu Pro Gly Gln Leu
305 310 315
Val Gly Arg Ile Ile Gln Lys Asp Pro Leu Arg Arg Phe Asp Gly
320 325 330
Tyr Leu Asn Gln Gly Ala Asn Asn Lys Lys Ile Ala Lys Asp Val

335 340 345
Phe Lys Lys Gly Asp Gln Ala Tyr Leu Thr Gly Asp Val Leu Val
350 355 360
Met Asp Glu Leu Gly Tyr Leu Tyr Phe Arg Asp Arg Thr Gly Asp
365 370 375
Thr Phe Arg Trp Lys Gly Glu Asn Val Ser Thr Thr Glu Val Glu
380 385 390
Gly Thr Leu Ser Arg Leu Leu Asp Met Ala Asp Val Ala Val Tyr
395 400 405
Gly Val Glu Val Pro Gly Thr Glu Gly Arg Ala Gly Met Ala Ala
410 415 420
Val Ala Ser Pro Thr Gly Asn Cys Asp Leu Glu Arg Phe Ala Gln
425 430 435
Val Leu Glu Lys Glu Leu Pro Leu Tyr Ala Arg Pro Ile Phe Leu
440 445 450
Arg Leu Leu Pro Glu Leu His Lys Thr Gly Thr Tyr Lys Phe Gln
455 460 465
Lys Thr Glu Leu Arg Lys Glu Gly Phe Asp Pro Ala Ile Val Lys
470 475 480
Asp Pro Leu Phe Tyr Leu Asp Ala Gln Lys Gly Arg Tyr Val Pro
485 490 495
Leu Asp Gln Glu Ala Tyr Ser Arg Ile Gln Ala Gly Glu Glu Lys
500 505 510
Leu

<210> 4
<211> 718
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2274290CD1

<400> 4
Met Leu Val His Leu Phe Arg Val Gly Ile Arg Gly Gly Pro Phe
1 5 10 15
Pro Gly Arg Leu Leu Pro Pro Leu Arg Phe Gln Thr Phe Ser Ala
20 25 30
Val Arg Tyr Ser Asp Gly Tyr Arg Ser Ser Ser Leu Leu Arg Ala
35 40 45
Val Ala His Leu Arg Ser Gln Leu Trp Ala His Leu Pro Arg Ala
50 55 60
Pro Leu Ala Pro Arg Trp Ser Pro Ser Ala Trp Cys Trp Val Gly
65 70 75
Gly Ala Leu Leu Gly Pro Met Val Leu Ser Lys His Pro His Leu
80 85 90
Cys Leu Val Ala Leu Cys Glu Ala Glu Glu Ala Pro Pro Ala Ser
95 100 105
Ser Thr Pro His Val Val Gly Ser Arg Phe Asn Trp Lys Leu Phe
110 115 120
Trp Gln Phe Leu His Pro His Leu Leu Val Leu Gly Val Ala Val
125 130 135
Val Leu Ala Leu Gly Ala Ala Leu Val Asn Val Gln Ile Pro Leu

```
                140                 145                 150
Leu Leu Gly Gln Leu Val Glu Val Val Ala Lys Tyr Thr Arg Asp
                155                 160                 165
His Val Gly Ser Phe Met Thr Glu Ser Gln Asn Leu Ser Thr His
                170                 175                 180
Leu Leu Ile Leu Tyr Gly Val Gln Gly Leu Leu Thr Phe Gly Tyr
                185                 190                 195
Leu Val Leu Leu Ser His Val Gly Glu Arg Met Ala Val Asp Met
                200                 205                 210
Arg Arg Ala Leu Phe Ser Ser Leu Leu Arg Gln Asp Ile Thr Phe
                215                 220                 225
Phe Asp Ala Asn Lys Thr Gly Gln Leu Val Ser Arg Leu Thr Thr
                230                 235                 240
Asp Val Gln Glu Phe Lys Ser Ser Phe Lys Leu Val Ile Ser Gln
                245                 250                 255
Gly Leu Arg Ser Cys Thr Gln Val Ala Gly Cys Leu Val Ser Leu
                260                 265                 270
Ser Met Leu Ser Thr Arg Leu Thr Leu Leu Leu Met Val Ala Thr
                275                 280                 285
Pro Ala Leu Met Gly Val Gly Thr Leu Met Gly Ser Gly Leu Arg
                290                 295                 300
Lys Leu Ser Arg Gln Cys Gln Glu Gln Ile Ala Arg Ala Met Gly
                305                 310                 315
Val Ala Asp Glu Ala Leu Gly Asn Val Arg Thr Val Arg Ala Phe
                320                 325                 330
Ala Met Glu Gln Arg Glu Glu Glu Arg Tyr Gly Ala Glu Leu Glu
                335                 340                 345
Ala Cys Arg Cys Arg Ala Glu Glu Leu Gly Arg Gly Ile Ala Leu
                350                 355                 360
Phe Gln Gly Leu Ser Asn Ile Ala Phe Asn Cys Met Val Leu Gly
                365                 370                 375
Thr Leu Phe Ile Gly Gly Ser Leu Val Ala Gly Gln Gln Leu Thr
                380                 385                 390
Gly Gly Asp Leu Met Ser Phe Leu Val Ala Ser Gln Thr Val Gln
                395                 400                 405
Arg Ser Met Ala Asn Leu Ser Val Leu Phe Gly Gln Val Val Arg
                410                 415                 420
Gly Leu Ser Ala Gly Ala Arg Val Phe Glu Tyr Met Ala Leu Asn
                425                 430                 435
Pro Cys Ile Pro Leu Ser Gly Gly Cys Cys Val Pro Lys Glu Gln
                440                 445                 450
Leu Arg Gly Ser Val Thr Phe Gln Asn Val Cys Phe Ser Tyr Pro
                455                 460                 465
Cys Arg Pro Gly Phe Glu Val Leu Lys Asp Phe Thr Leu Thr Leu
                470                 475                 480
Pro Pro Gly Lys Ile Val Ala Leu Val Gly Gln Ser Gly Gly Gly
                485                 490                 495
Lys Thr Thr Val Ala Ser Leu Leu Glu Arg Phe Tyr Asp Pro Thr
                500                 505                 510
Ala Gly Val Val Met Leu Asp Gly Arg Asp Leu Arg Thr Leu Asp
                515                 520                 525
Pro Ser Trp Leu Arg Gly Gln Val Val Gly Phe Ile Ser Gln Glu
                530                 535                 540
Pro Val Leu Phe Gly Thr Thr Ile Met Glu Asn Ile Arg Phe Gly
                545                 550                 555
```

```
Lys Leu Glu Ala Ser Asp Glu Glu Val Tyr Thr Ala Ala Arg Glu
                560                 565                 570
Ala Asn Ala His Glu Phe Ile Thr Ser Phe Pro Glu Gly Tyr Asn
                575                 580                 585
Thr Val Val Gly Glu Arg Gly Thr Thr Leu Ser Gly Gly Gln Lys
                590                 595                 600
Gln Arg Leu Ala Ile Ala Arg Ala Leu Ile Lys Gln Pro Thr Val
                605                 610                 615
Leu Ile Leu Asp Glu Ala Thr Ser Ala Leu Asp Ala Glu Ser Glu
                620                 625                 630
Arg Val Val Gln Glu Ala Leu Asp Arg Ala Ser Ala Gly Arg Thr
                635                 640                 645
Val Leu Val Ile Ala His Arg Leu Ser Thr Val Arg Gly Ala His
                650                 655                 660
Cys Ile Val Val Met Ala Asp Gly Arg Val Trp Glu Ala Gly Thr
                665                 670                 675
His Glu Glu Leu Leu Lys Lys Gly Gly Leu Tyr Ala Glu Leu Ile
                680                 685                 690
Arg Arg Gln Ala Leu Asp Ala Pro Arg Thr Ala Ala Pro Pro Pro
                695                 700                 705
Lys Lys Pro Glu Gly Pro Arg Ser His Gln His Lys Ser
                710                 715
```

<210> 5
<211> 635
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2740029CD1

<400> 5

```
Met Ser Val Gly Val Ser Thr Ser Ala Pro Leu Ser Pro Thr Ser
  1               5                  10                  15
Gly Thr Ser Val Gly Met Ser Thr Phe Ser Ile Met Asp Tyr Val
                 20                  25                  30
Val Phe Val Leu Leu Leu Val Leu Ser Leu Ala Ile Gly Leu Tyr
                 35                  40                  45
His Ala Cys Arg Gly Trp Gly Arg His Thr Val Gly Glu Leu Leu
                 50                  55                  60
Met Ala Asp Arg Lys Met Gly Cys Leu Pro Val Ala Leu Ser Leu
                 65                  70                  75
Leu Ala Thr Phe Gln Ser Ala Val Ala Ile Leu Gly Val Pro Ser
                 80                  85                  90
Glu Ile Tyr Arg Phe Gly Thr Gln Tyr Trp Phe Leu Gly Cys Cys
                 95                 100                 105
Tyr Phe Leu Gly Leu Leu Ile Pro Ala His Ile Phe Ile Pro Val
                110                 115                 120
Phe Tyr Arg Leu His Leu Thr Ser Ala Tyr Glu Tyr Leu Glu Leu
                125                 130                 135
Arg Phe Asn Lys Thr Val Arg Val Cys Gly Thr Val Thr Phe Ile
                140                 145                 150
Phe Gln Met Val Ile Tyr Met Gly Val Val Leu Tyr Ala Pro Ser
```

155 160 165
Leu Ala Leu Asn Ala Val Thr Gly Phe Asp Leu Trp Leu Ser Val
170 175 180
Leu Ala Leu Gly Ile Val Cys Thr Val Tyr Thr Ala Leu Gly Gly
185 190 195
Leu Lys Ala Val Ile Trp Thr Asp Val Phe Gln Thr Leu Val Met
200 205 210
Phe Leu Gly Gln Leu Ala Val Ile Ile Val Gly Ser Ala Lys Val
215 220 225
Gly Gly Leu Gly Arg Val Trp Ala Val Ala Ser Gln His Gly Arg
230 235 240
Ile Ser Gly Phe Glu Leu Asp Pro Asp Pro Phe Val Arg His Thr
245 250 255
Phe Trp Thr Leu Ala Phe Gly Gly Val Phe Met Met Leu Ser Leu
260 265 270
Tyr Gly Val Asn Gln Ala Gln Val Gln Arg Tyr Leu Ser Ser Arg
275 280 285
Thr Glu Lys Ala Ala Val Leu Ser Cys Tyr Ala Val Phe Pro Phe
290 295 300
Gln Gln Val Ser Leu Cys Val Gly Cys Leu Ile Gly Leu Val Met
305 310 315
Phe Ala Tyr Tyr Gln Glu Tyr Pro Met Ser Ile Gln Gln Ala Gln
320 325 330
Ala Ala Pro Asp Gln Phe Val Leu Tyr Phe Val Met Asp Leu Leu
335 340 345
Lys Gly Leu Pro Gly Leu Pro Gly Leu Phe Ile Ala Cys Leu Phe
350 355 360
Ser Gly Ser Leu Ser Thr Ile Ser Ser Ala Phe Asn Ser Leu Ala
365 370 375
Thr Val Thr Met Glu Asp Leu Ile Arg Pro Trp Phe Pro Glu Phe
380 385 390
Ser Glu Ala Arg Ala Ile Met Leu Ser Arg Gly Leu Ala Phe Gly
395 400 405
Tyr Gly Leu Leu Cys Leu Gly Met Ala Tyr Ile Ser Ser Gln Met
410 415 420
Gly Pro Val Leu Gln Ala Ala Ile Ser Ile Phe Gly Met Val Gly
425 430 435
Gly Pro Leu Leu Gly Leu Phe Cys Leu Gly Met Phe Phe Pro Cys
440 445 450
Ala Asn Pro Pro Gly Ala Val Val Gly Leu Leu Ala Gly Leu Val
455 460 465
Met Ala Phe Trp Ile Gly Ile Gly Ser Ile Val Thr Ser Met Gly
470 475 480
Ser Ser Met Pro Pro Ser Pro Ser Asn Gly Ser Ser Phe Ser Leu
485 490 495
Pro Thr Asn Leu Thr Val Ala Thr Val Thr Thr Leu Met Pro Leu
500 505 510
Thr Thr Phe Ser Lys Pro Thr Gly Leu Gln Arg Phe Tyr Ser Leu
515 520 525
Ser Tyr Leu Trp Tyr Ser Ala His Asn Ser Thr Thr Val Ile Val
530 535 540
Val Gly Leu Ile Val Ser Leu Leu Thr Gly Arg Met Arg Gly Arg
545 550 555
Ser Leu Asn Pro Ala Thr Ile Tyr Pro Val Leu Pro Lys Leu Leu
560 565 570

```
Ser Leu Leu Pro Leu Ser Cys Gln Lys Arg Leu His Cys Arg Ser
                575                 580                 585
Tyr Gly Gln Asp His Leu Asp Thr Gly Leu Phe Pro Glu Lys Pro
                590                 595                 600
Arg Asn Gly Val Leu Gly Asp Ser Arg Asp Lys Glu Ala Met Ala
                605                 610                 615
Leu Asp Gly Thr Ala Tyr Gln Gly Ser Ser Ser Thr Cys Ile Leu
                620                 625                 630
Gln Glu Thr Ser Leu
                635
```

<210> 6
<211> 535
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2414415CD1

<400> 6

```
Met Glu Glu Gly Ala Arg His Arg Asn Asn Thr Glu Lys Lys His
  1                 5                  10                  15
Pro Gly Gly Gly Glu Ser Asp Ala Ser Pro Glu Ala Gly Ser Gly
                 20                  25                  30
Gly Gly Gly Val Ala Leu Lys Lys Glu Ile Gly Leu Val Ser Ala
                 35                  40                  45
Cys Gly Ile Ile Val Gly Asn Ile Ile Gly Ser Gly Ile Phe Val
                 50                  55                  60
Ser Pro Lys Gly Val Leu Glu Asn Ala Gly Ser Val Gly Leu Ala
                 65                  70                  75
Leu Ile Val Trp Ile Val Thr Gly Phe Ile Thr Val Val Gly Ala
                 80                  85                  90
Leu Cys Tyr Ala Glu Leu Gly Val Thr Ile Pro Lys Ser Gly Gly
                 95                 100                 105
Asp Tyr Ser Tyr Val Lys Asp Ile Phe Gly Gly Leu Ala Gly Phe
                110                 115                 120
Leu Arg Leu Trp Ile Ala Val Leu Val Ile Tyr Pro Thr Asn Gln
                125                 130                 135
Ala Val Ile Ala Leu Thr Phe Ser Asn Tyr Val Leu Gln Pro Leu
                140                 145                 150
Phe Pro Thr Cys Phe Pro Pro Glu Ser Gly Leu Arg Leu Leu Ala
                155                 160                 165
Ala Ile Cys Leu Leu Leu Leu Thr Trp Val Asn Cys Ser Ser Val
                170                 175                 180
Arg Trp Ala Thr Arg Val Gln Asp Ile Phe Thr Ala Gly Lys Leu
                185                 190                 195
Leu Ala Leu Ala Leu Ile Ile Ile Met Gly Ile Val Gln Ile Cys
                200                 205                 210
Lys Gly Glu Tyr Phe Trp Leu Glu Pro Lys Asn Ala Phe Glu Asn
                215                 220                 225
Phe Gln Glu Pro Asp Ile Gly Leu Val Ala Leu Ala Phe Leu Gln
                230                 235                 240
Gly Ser Phe Ala Tyr Gly Gly Trp Asn Phe Leu Asn Tyr Val Thr
```

```
                245                 250                 255
Glu Glu Leu Val Asp Pro Tyr Lys Asn Leu Pro Arg Ala Ile Phe
                260                 265                 270
Ile Ser Ile Pro Leu Val Thr Phe Val Tyr Val Phe Ala Asn Val
                275                 280                 285
Ala Tyr Val Thr Ala Met Ser Pro Gln Glu Leu Leu Ala Ser Asn
                290                 295                 300
Ala Val Ala Val Thr Phe Gly Glu Lys Leu Leu Gly Val Met Ala
                305                 310                 315
Trp Ile Met Pro Ile Ser Val Ala Leu Ser Thr Phe Gly Gly Val
                320                 325                 330
Asn Gly Ser Leu Phe Thr Ser Ser Arg Leu Phe Phe Ala Gly Ala
                335                 340                 345
Arg Glu Gly His Leu Pro Ser Val Leu Ala Met Ile His Val Lys
                350                 355                 360
Arg Cys Thr Pro Ile Pro Ala Leu Leu Phe Thr Cys Ile Ser Thr
                365                 370                 375
Leu Leu Met Leu Val Thr Ser Asp Met Tyr Thr Leu Ile Asn Tyr
                380                 385                 390
Val Gly Phe Ile Asn Tyr Leu Phe Tyr Gly Val Thr Val Ala Gly
                395                 400                 405
Gln Ile Val Leu Arg Trp Lys Lys Pro Asp Ile Pro Arg Pro Ile
                410                 415                 420
Lys Ile Asn Leu Leu Phe Pro Ile Ile Tyr Leu Leu Phe Trp Ala
                425                 430                 435
Phe Leu Leu Val Phe Ser Leu Trp Ser Glu Pro Val Val Cys Gly
                440                 445                 450
Ile Gly Leu Ala Ile Met Leu Thr Gly Val Pro Val Tyr Phe Leu
                455                 460                 465
Gly Val Tyr Trp Gln His Lys Pro Lys Cys Phe Ser Asp Phe Ile
                470                 475                 480
Glu Leu Leu Thr Leu Val Ser Gln Lys Met Cys Val Val Val Tyr
                485                 490                 495
Pro Glu Val Glu Arg Gly Ser Gly Thr Glu Glu Ala Asn Glu Asp
                500                 505                 510
Met Glu Glu Gln Gln Gln Pro Met Tyr Gln Pro Thr Pro Thr Lys
                515                 520                 525
Asp Lys Asp Val Ala Gly Gln Pro Gln Pro
                530                 535
```

<210> 7
<211> 456
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2466714CD1

```
<400> 7
Met Glu Ala Ser Trp Gly Ser Phe Asn Ala Glu Arg Gly Trp Tyr
  1               5                  10                  15
Val Ser Val Gln Gln Pro Glu Glu Ala Glu Ala Glu Glu Leu Ser
                20                  25                  30
```

Pro Leu Leu Ser Asn Glu Leu His Arg Gln Arg Ser Pro Gly Val
35 40 45
Ser Phe Gly Leu Ser Val Phe Asn Leu Met Asn Ala Ile Met Gly
50 55 60
Ser Gly Ile Leu Gly Leu Ala Tyr Val Met Ala Asn Thr Gly Val
65 70 75
Phe Gly Phe Ser Phe Leu Leu Leu Thr Val Ala Leu Leu Ala Ser
80 85 90
Tyr Ser Val His Leu Leu Leu Ser Met Cys Ile Gln Thr Ala Val
95 100 105
Thr Ser Tyr Glu Asp Leu Gly Leu Phe Ala Phe Gly Leu Pro Gly
110 115 120
Lys Leu Val Val Ala Gly Thr Ile Ile Ile Gln Asn Ile Gly Ala
125 130 135
Met Ser Ser Tyr Leu Leu Ile Ile Lys Thr Glu Leu Pro Ala Ala
140 145 150
Ile Ala Glu Phe Leu Thr Gly Asp Tyr Asn Arg Tyr Trp Tyr Leu
155 160 165
Asp Gly Gln Thr Leu Leu Ile Ile Ile Cys Val Gly Ile Val Phe
170 175 180
Pro Leu Ala Leu Leu Pro Lys Ile Gly Phe Leu Gly Tyr Thr Ser
185 190 195
Ser Leu Ser Phe Phe Phe Met Met Phe Phe Ala Leu Val Val Ile
200 205 210
Ile Lys Lys Trp Ser Ile Pro Cys Pro Leu Thr Leu Asn Tyr Val
215 220 225
Glu Lys Gly Phe Gln Ile Ser Asn Val Thr Asp Asp Cys Lys Pro
230 235 240
Lys Leu Phe His Phe Ser Lys Glu Ser Ala Tyr Ala Leu Pro Thr
245 250 255
Met Ala Phe Ser Phe Leu Cys His Thr Ser Ile Leu Pro Ile Tyr
260 265 270
Cys Glu Leu Gln Ser Pro Ser Lys Lys Arg Met Gln Asn Val Thr
275 280 285
Asn Thr Ala Ile Ala Leu Ser Phe Leu Ile Tyr Phe Ile Ser Ala
290 295 300
Leu Phe Gly Tyr Leu Thr Phe Tyr Asp Lys Val Glu Ser Glu Leu
305 310 315
Leu Lys Gly Tyr Ser Lys Tyr Leu Ser His Asp Val Val Val Met
320 325 330
Thr Val Lys Leu Cys Ile Leu Phe Ala Val Leu Leu Thr Val Pro
335 340 345
Leu Ile His Phe Pro Ala Arg Lys Ala Val Thr Met Met Phe Phe
350 355 360
Ser Asn Phe Pro Phe Ser Trp Ile Arg His Phe Leu Ile Thr Leu
365 370 375
Ala Leu Asn Ile Ile Ile Val Leu Leu Ala Ile Tyr Val Pro Asp
380 385 390
Ile Arg Asn Val Phe Gly Val Val Gly Ala Ser Thr Ser Thr Cys
395 400 405
Leu Ile Phe Ile Phe Pro Gly Leu Phe Tyr Leu Lys Leu Ser Arg
410 415 420
Glu Asp Phe Leu Ser Trp Lys Lys Leu Gly Ala Phe Val Leu Leu
425 430 435
Ile Phe Gly Ile Leu Val Gly Asn Phe Ser Leu Ala Leu Ile Ile

440 445 450
Phe Asp Trp Ile Asn Lys
455

<210> 8
<211> 325
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2617942CD1

<400> 8
Met Phe Ala Asn Leu Lys Tyr Val Ser Leu Gly Ile Leu Val Phe
1 5 10 15
Gln Thr Thr Ser Leu Val Leu Thr Met Arg Tyr Ser Arg Thr Leu
20 25 30
Lys Glu Glu Gly Pro Arg Tyr Leu Ser Ser Thr Ala Val Val Val
35 40 45
Ala Glu Leu Leu Lys Ile Met Ala Cys Ile Leu Leu Val Tyr Lys
50 55 60
Asp Ser Lys Cys Ser Leu Arg Ala Leu Asn Arg Val Leu His Asp
65 70 75
Glu Ile Leu Asn Lys Pro Met Glu Thr Leu Lys Leu Ala Ile Pro
80 85 90
Ser Gly Ile Tyr Thr Leu Gln Asn Asn Leu Leu Tyr Val Ala Leu
95 100 105
Ser Asn Leu Asp Ala Ala Thr Tyr Gln Val Thr Tyr Gln Leu Lys
110 115 120
Ile Leu Thr Thr Ala Leu Phe Ser Val Ser Met Leu Ser Lys Lys
125 130 135
Leu Gly Val Tyr Gln Trp Leu Ser Leu Val Ile Leu Met Thr Gly
140 145 150
Val Ala Phe Val Gln Trp Pro Ser Asp Ser Gln Leu Asp Ser Lys
155 160 165
Glu Leu Ser Ala Gly Ser Gln Phe Val Gly Leu Met Ala Val Leu
170 175 180
Thr Ala Cys Phe Ser Ser Gly Phe Ala Gly Val Tyr Phe Glu Lys
185 190 195
Ile Leu Lys Glu Thr Lys Gln Ser Val Trp Ile Arg Asn Ile Gln
200 205 210
Leu Gly Phe Phe Gly Ser Ile Phe Gly Leu Met Gly Val Tyr Ile
215 220 225
Tyr Asp Gly Glu Leu Val Ser Lys Asn Gly Phe Phe Gln Gly Tyr
230 235 240
Asn Arg Leu Thr Trp Ile Val Val Val Leu Gln Ala Leu Gly Gly
245 250 255
Leu Val Ile Ala Ala Val Ile Lys Tyr Ala Asp Asn Ile Leu Lys
260 265 270
Gly Phe Ala Thr Ser Leu Ser Ile Ile Leu Ser Thr Leu Ile Ser
275 280 285
Tyr Phe Trp Leu Gln Asp Phe Val Pro Thr Ser Val Phe Phe Leu
290 295 300

```
Gly Ala Ile Leu Val Ile Thr Ala Thr Phe Leu Tyr Gly Tyr Asp
                305                 310                 315
Pro Lys Pro Ala Gly Asn Pro Thr Lys Ala
                320                 325
```

<210> 9
<211> 178
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2945431CD1

<400> 9
```
Met Ser Leu Ser Pro Arg Ser Gln Leu Ala Ile Ile Pro Gln Glu
  1                 5                  10                  15
Pro Phe Leu Phe Ser Gly Thr Val Arg Glu Asn Leu Asp Pro Gln
                 20                  25                  30
Gly Leu His Lys Asp Arg Ala Leu Trp Gln Ala Leu Lys Gln Cys
                 35                  40                  45
His Leu Ser Glu Val Ile Thr Ser Met Gly Gly Leu Asp Gly Glu
                 50                  55                  60
Leu Gly Glu Gly Gly Arg Ser Leu Ser Leu Gly Gln Arg Gln Leu
                 65                  70                  75
Leu Cys Leu Ala Arg Ala Leu Leu Thr Asp Ala Lys Ile Leu Cys
                 80                  85                  90
Ile Asp Glu Ala Thr Ala Ser Val Asp Gln Lys Thr Asp Gln Leu
                 95                 100                 105
Leu Gln Gln Thr Ile Cys Lys Arg Phe Ala Asn Lys Thr Val Leu
                110                 115                 120
Thr Ile Ala His Arg Leu Asn Thr Ile Leu Asn Ser Asp Arg Val
                125                 130                 135
Leu Val Leu Gln Ala Gly Arg Val Val Glu Leu Asp Ser Pro Ala
                140                 145                 150
Thr Leu Arg Asn Gln Pro His Ser Leu Phe Gln Gln Leu Leu Gln
                155                 160                 165
Ser Ser Gln Gln Gly Val Pro Ala Ser Leu Gly Gly Pro
                170                 175
```

<210> 10
<211> 255
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4074113CD1

<400> 10
```
Met Glu Arg Glu Met Glu Gly Arg Pro Leu His Asn Glu Gly Trp
  1                 5                  10                  15
Ile Asp Arg Ser Arg Val Gln Gln Lys Asp Leu Pro Asn Lys Cys
```

```
                  20                  25                  30
Pro Gln Thr Leu Trp Ser Glu Gln Ala Phe Pro Pro Asn Pro Gly
                  35                  40                  45
Gln Val Gly Ile Val Gly Arg Thr Gly Ala Gly Lys Ser Ser Leu
                  50                  55                  60
Ala Ser Gly Leu Leu Arg Leu Pro Glu Ala Ala Glu Gly Gly Ile
                  65                  70                  75
Trp Ile Asp Gly Val Pro Ile Ala His Val Gly Leu His Thr Leu
                  80                  85                  90
Arg Ser Arg Ile Ser Ile Ile Pro Gln Asp Pro Ile Leu Phe Pro
                  95                 100                 105
Gly Ser Leu Arg Met Asn Leu Asp Leu Leu Gln Glu His Ser Asp
                 110                 115                 120
Glu Ala Ile Trp Ala Ala Leu Glu Thr Val Gln Leu Lys Ala Leu
                 125                 130                 135
Val Ala Ser Leu Pro Gly Gln Leu Gln Tyr Lys Cys Ala Asp Arg
                 140                 145                 150
Gly Glu Asp Leu Ser Val Gly Gln Lys Gln Leu Leu Cys Leu Ala
                 155                 160                 165
Arg Ala Leu Leu Arg Lys Thr Gln Ile Leu Ile Leu Asp Glu Ala
                 170                 175                 180
Thr Ala Ala Val Asp Pro Gly Thr Glu Leu Gln Met Gln Ala Met
                 185                 190                 195
Leu Gly Ser Trp Phe Ala Gln Cys Thr Val Leu Leu Ile Ala His
                 200                 205                 210
Arg Leu Arg Ser Val Met Asp Cys Ala Arg Val Leu Val Met Asp
                 215                 220                 225
Lys Gly Gln Val Ala Glu Ser Gly Ser Pro Ala Gln Leu Leu Ala
                 230                 235                 240
Gln Lys Gly Leu Phe Tyr Arg Leu Ala Gln Glu Ser Gly Leu Val
                 245                 250                 255
```

<210> 11
<211> 462
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1413743CD1

<400> 11

```
Met Ala Gln Val Ser Ile Asn Asn Asp Tyr Ser Glu Trp Asp Leu
  1               5                  10                  15
Ser Thr Asp Ala Gly Glu Arg Ala Arg Leu Leu Gln Ser Pro Cys
                 20                  25                  30
Val Asp Thr Ala Pro Lys Ser Glu Trp Glu Ala Ser Pro Gly Gly
                 35                  40                  45
Leu Asp Arg Gly Thr Thr Ser Thr Leu Gly Ala Ile Phe Ile Val
                 50                  55                  60
Val Asn Ala Cys Leu Gly Ala Gly Leu Leu Asn Phe Pro Ala Ala
                 65                  70                  75
Phe Ser Thr Ala Gly Gly Val Ala Ala Gly Ile Ala Leu Gln Met
                 80                  85                  90
```

Gly Met Leu Val Phe Ile Ile Ser Gly Leu Val Ile Leu Ala Tyr
95 100 105
Cys Ser Gln Ala Ser Asn Glu Arg Thr Tyr Gln Glu Val Val Trp
110 115 120
Ala Val Cys Gly Lys Leu Thr Gly Val Leu Cys Glu Val Ala Ile
125 130 135
Ala Val Tyr Thr Phe Gly Thr Cys Ile Ala Phe Leu Ile Ile Ile
140 145 150
Gly Asp Gln Gln Asp Lys Ile Ile Ala Val Met Ala Lys Glu Pro
155 160 165
Glu Gly Ala Ser Gly Pro Trp Tyr Thr Asp Arg Lys Phe Thr Ile
170 175 180
Ser Leu Thr Ala Phe Leu Phe Ile Leu Pro Leu Ser Ile Pro Arg
185 190 195
Glu Ile Gly Phe Gln Lys Tyr Ala Ser Phe Leu Ser Val Val Gly
200 205 210
Thr Trp Tyr Val Thr Ala Ile Val Ile Ile Lys Tyr Ile Trp Pro
215 220 225
Asp Lys Glu Met Thr Pro Gly Asn Ile Leu Thr Arg Pro Ala Ser
230 235 240
Trp Met Ala Val Phe Asn Ala Met Pro Thr Ile Cys Phe Gly Phe
245 250 255
Gln Cys His Val Ser Ser Val Pro Val Phe Asn Ser Met Gln Gln
260 265 270
Pro Glu Val Lys Thr Trp Gly Gly Val Val Thr Ala Ala Met Val
275 280 285
Ile Ala Leu Ala Val Tyr Met Gly Thr Gly Ile Cys Gly Phe Leu
290 295 300
Thr Phe Gly Ala Ala Val Asp Pro Asp Val Leu Leu Ser Tyr Pro
305 310 315
Ser Glu Asp Met Ala Val Ala Val Ala Arg Ala Phe Ile Ile Leu
320 325 330
Ser Val Leu Thr Ser Tyr Pro Ile Leu His Phe Cys Gly Arg Ala
335 340 345
Val Val Glu Gly Leu Trp Leu Arg Tyr Gln Gly Val Pro Val Glu
350 355 360
Glu Asp Val Gly Arg Glu Arg Arg Arg Arg Val Leu Gln Thr Leu
365 370 375
Val Trp Phe Leu Leu Thr Leu Leu Leu Ala Leu Phe Ile Pro Asp
380 385 390
Ile Gly Lys Val Ile Ser Val Ile Gly Gly Leu Ala Ala Cys Phe
395 400 405
Ile Phe Val Phe Pro Gly Leu Cys Leu Ile Gln Ala Lys Leu Ser
410 415 420
Glu Met Glu Glu Val Lys Pro Ala Ser Trp Trp Val Leu Val Ser
425 430 435
Tyr Gly Val Leu Leu Val Thr Leu Gly Ala Phe Ile Phe Gly Gln
440 445 450
Thr Thr Ala Asn Ala Ile Phe Val Asp Leu Leu Ala
455 460

<210> 12
<211> 758
<212> PRT

<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1733477CD1

<400> 12

```
Met Gly Leu Ala Asp Ala Ser Gly Pro Arg Asp Thr Gln Ala Leu
  1               5                  10                  15
Leu Ser Ala Thr Gln Ala Met Asp Leu Arg Arg Arg Asp Tyr His
             20                  25                  30
Met Glu Arg Pro Leu Leu Asn Gln Glu His Leu Glu Glu Leu Gly
         35                  40                  45
Arg Trp Gly Ser Ala Pro Arg Thr His Gln Trp Arg Thr Trp Leu
     50                  55                  60
Gln Cys Ser Arg Ala Arg Ala Tyr Ala Leu Leu Leu Gln His Leu
 65                  70                  75
Pro Val Leu Val Trp Leu Pro Arg Tyr Pro Val Arg Asp Trp Leu
             80                  85                  90
Leu Gly Asp Leu Leu Ser Gly Leu Ser Val Ala Ile Met Gln Leu
         95                 100                 105
Pro Gln Gly Leu Ala Tyr Ala Leu Leu Ala Gly Leu Pro Pro Val
        110                 115                 120
Phe Gly Leu Tyr Ser Ser Phe Tyr Pro Val Phe Ile Tyr Phe Leu
        125                 130                 135
Phe Gly Thr Ser Arg His Ile Ser Val Gly Thr Phe Ala Val Met
        140                 145                 150
Ser Val Met Val Gly Gly Val Thr Glu Ser Leu Ala Pro Gln Ala
        155                 160                 165
Leu Asn Asp Ser Met Ile Asn Glu Thr Ala Arg Asp Ala Ala Arg
        170                 175                 180
Val Gln Val Ala Ser Thr Leu Ser Val Leu Val Gly Leu Phe Gln
        185                 190                 195
Val Gly Leu Gly Leu Ile His Phe Gly Phe Val Val Thr Tyr Leu
        200                 205                 210
Ser Glu Pro Leu Val Arg Gly Tyr Thr Thr Ala Ala Ala Val Gln
        215                 220                 225
Val Phe Val Ser Gln Leu Lys Tyr Val Phe Gly Leu His Leu Ser
        230                 235                 240
Ser His Ser Gly Pro Leu Ser Leu Ile Tyr Thr Val Leu Glu Val
        245                 250                 255
Cys Trp Lys Leu Pro Gln Ser Lys Val Gly Thr Val Val Thr Ala
        260                 265                 270
Ala Val Ala Gly Val Val Leu Val Val Val Lys Leu Leu Asn Asp
        275                 280                 285
Lys Leu Gln Gln Gln Leu Pro Met Pro Ile Pro Gly Glu Leu Leu
        290                 295                 300
Thr Leu Ile Gly Ala Thr Gly Ile Ser Tyr Gly Met Gly Leu Lys
        305                 310                 315
His Arg Phe Glu Val Asp Val Val Gly Asn Ile Pro Ala Gly Leu
        320                 325                 330
Val Pro Pro Val Ala Pro Asn Thr Gln Leu Phe Ser Lys Leu Val
        335                 340                 345
Gly Ser Ala Phe Thr Ile Ala Val Val Gly Phe Ala Ile Ala Ile
        350                 355                 360
```

```
Ser Leu Gly Lys Ile Phe Ala Leu Arg His Gly Tyr Arg Val Asp
                365                 370                 375
Ser Asn Gln Glu Leu Val Ala Leu Gly Leu Ser Asn Leu Ile Gly
                380                 385                 390
Gly Ile Phe Gln Cys Phe Pro Val Ser Cys Ser Met Ser Arg Ser
                395                 400                 405
Leu Val Gln Glu Ser Thr Gly Gly Asn Ser Gln Val Ala Gly Ala
                410                 415                 420
Ile Ser Ser Leu Phe Ile Leu Leu Ile Ile Val Lys Leu Gly Glu
                425                 430                 435
Leu Phe His Asp Leu Pro Lys Ala Val Leu Ala Ala Ile Ile Ile
                440                 445                 450
Val Asn Leu Lys Gly Met Leu Arg Gln Leu Ser Asp Met Arg Ser
                455                 460                 465
Leu Trp Lys Ala Asn Arg Ala Asp Leu Leu Ile Trp Leu Val Thr
                470                 475                 480
Phe Thr Ala Thr Ile Leu Leu Asn Leu Asp Leu Gly Leu Val Val
                485                 490                 495
Ala Val Ile Phe Ser Leu Leu Leu Val Val Val Arg Thr Gln Met
                500                 505                 510
Pro His Tyr Ser Val Leu Gly Gln Val Pro Asp Thr Asp Ile Tyr
                515                 520                 525
Arg Asp Val Ala Glu Tyr Ser Glu Ala Lys Glu Val Arg Gly Val
                530                 535                 540
Lys Val Phe Arg Ser Ser Ala Thr Val Tyr Phe Ala Asn Ala Glu
                545                 550                 555
Phe Tyr Ser Asp Ala Leu Lys Gln Arg Cys Gly Val Asp Val Asp
                560                 565                 570
Phe Leu Ile Ser Gln Lys Lys Lys Leu Leu Lys Lys Gln Glu Gln
                575                 580                 585
Leu Lys Leu Lys Gln Leu Gln Lys Glu Glu Lys Leu Arg Lys Gln
                590                 595                 600
Ala Ala Ser Pro Lys Gly Ala Ser Val Ser Ile Asn Val Asn Thr
                605                 610                 615
Ser Leu Glu Asp Met Arg Ser Asn Asn Val Glu Asp Cys Lys Met
                620                 625                 630
Met Val Ser Ser Gly Asp Lys Met Glu Asp Ala Thr Ala Asn Gly
                635                 640                 645
Gln Glu Asp Ser Lys Ala Pro Asp Gly Ser Thr Leu Lys Ala Leu
                650                 655                 660
Gly Leu Pro Gln Pro Asp Phe His Ser Leu Ile Leu Asp Leu Gly
                665                 670                 675
Ala Leu Ser Phe Val Asp Thr Val Cys Leu Lys Ser Leu Lys Asn
                680                 685                 690
Ile Phe His Asp Phe Arg Glu Ile Glu Val Glu Val Tyr Met Ala
                695                 700                 705
Ala Cys His Ser Pro Val Val Ser Gln Leu Glu Ala Gly His Phe
                710                 715                 720
Phe Asp Ala Ser Ile Thr Lys Lys His Leu Phe Ala Ser Val His
                725                 730                 735
Asp Ala Val Thr Phe Ala Leu Gln His Pro Arg Pro Val Pro Asp
                740                 745                 750
Ser Pro Val Ser Val Thr Arg Leu
                755
```

<210> 13
<211> 336
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2641908CD1

<400> 13
Met Met Gly Pro Gly Leu Ala Phe Gly Leu Gly Ser Leu Met Leu
1 5 10 15
Arg Leu Tyr Val Asp Ile Asn Gln Met Pro Glu Gly Gly Ile Ser
20 25 30
Leu Thr Ile Lys Asp Pro Arg Trp Val Gly Ala Trp Trp Leu Gly
35 40 45
Phe Leu Ile Ala Ala Gly Ala Val Ala Leu Ala Ala Ile Pro Tyr
50 55 60
Phe Phe Phe Pro Lys Glu Met Pro Lys Glu Lys Arg Glu Leu Gln
65 70 75
Phe Arg Arg Lys Val Leu Ala Val Thr Asp Ser Pro Ala Arg Lys
80 85 90
Gly Lys Asp Ser Pro Ser Lys Gln Ser Pro Gly Glu Ser Thr Lys
95 100 105
Lys Gln Asp Gly Leu Val Gln Ile Ala Pro Asn Leu Thr Val Ile
110 115 120
Gln Phe Ile Lys Val Phe Pro Arg Val Leu Leu Gln Thr Leu Arg
125 130 135
His Pro Ile Phe Leu Leu Val Val Leu Ser Gln Val Cys Leu Ser
140 145 150
Ser Met Ala Ala Gly Met Ala Thr Phe Leu Pro Lys Phe Leu Glu
155 160 165
Arg Gln Phe Ser Ile Thr Ala Ser Tyr Ala Asn Leu Leu Ile Gly
170 175 180
Cys Leu Ser Phe Pro Ser Val Ile Val Gly Ile Val Val Gly Gly
185 190 195
Val Leu Val Lys Arg Leu His Leu Gly Pro Val Gly Cys Gly Ala
200 205 210
Leu Cys Leu Leu Gly Met Leu Leu Cys Leu Phe Phe Ser Leu Pro
215 220 225
Leu Phe Phe Ile Gly Cys Ser Ser His Gln Ile Ala Gly Ile Thr
230 235 240
His Gln Thr Ser Ala His Pro Gly Leu Glu Leu Ser Pro Ser Cys
245 250 255
Met Glu Ala Cys Ser Cys Pro Leu Asp Gly Phe Asn Pro Val Cys
260 265 270
Asp Pro Ser Thr Arg Val Glu Tyr Ile Thr Pro Cys His Ala Gly
275 280 285
Cys Ser Ser Trp Val Val Gln Asp Ala Leu Asp Asn Ser Gln Ser
290 295 300
Pro Pro Thr Ser His Pro His Ala Gly His Gln His Leu Asn Leu
305 310 315
Arg Leu Leu Gln Gly Glu Thr Trp Ala Ala Leu Ala Gly Ala Glu
320 325 330
Glu Pro Val Asp Gly Ala

335

<210> 14
<211> 103
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2656554CD1

<400> 14
Met Glu Arg Gln Ser Arg Val Met Ser Glu Lys Asp Glu Tyr Gln
1 5 10 15
Phe Gln His Gln Gly Ala Val Glu Leu Leu Val Phe Asn Phe Leu
20 25 30
Leu Ile Leu Thr Ile Leu Thr Ile Trp Leu Phe Lys Asn His Arg
35 40 45
Phe Arg Phe Leu His Glu Thr Gly Gly Ala Met Val Tyr Asp Lys
50 55 60
Pro Pro Lys Phe Ala Met Ser Arg Glu Gln Met Ser Gln Ser Cys
65 70 75
Ser His Thr Ala His Asn Ala Ser Leu Leu Thr Asp Ala Gly Pro
80 85 90
Leu Ser Cys Gly Glu Ser Arg Ala Ser Cys Leu Phe Leu
95 100

<210> 15
<211> 123
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2719228CD1

<400> 15
Met Gln Gly Met Gly Leu Gly Leu Ser Ser Val Phe Ala Leu Cys
1 5 10 15
Leu Gly His Thr Ser Ser Phe Cys Glu Ser Val Val Phe Ala Ser
20 25 30
Ala Ser Ile Gly Leu Gln Thr Phe Asn His Ser Gly Ile Ser Val
35 40 45
Asn Ile Gln Asp Leu Ala Pro Ser Cys Ala Gly Phe Leu Phe Gly
50 55 60
Val Ala Asn Thr Ala Gly Ala Leu Ala Gly Val Val Gly Val Cys
65 70 75
Leu Gly Gly Tyr Leu Met Glu Thr Thr Gly Ser Trp Thr Cys Leu
80 85 90
Phe Asn Leu Val Ala Ile Ile Ser Asn Leu Gly Leu Cys Thr Phe
95 100 105
Leu Val Phe Gly Gln Ala Gln Arg Val Asp Leu Ser Ser Thr His
110 115 120

Glu Asp Leu

<210> 16
<211> 222
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3657824CD1

<400> 16
Met Lys Gln Glu Ser Ala Ala Pro Asn Thr Pro Pro Thr Ser Gln
1 5 10 15
Ser Pro Thr Pro Ser Ala Gln Phe Pro Arg Asn Asp Gly Asp Pro
20 25 30
Gln Ala Leu Trp Ile Phe Gly Tyr Gly Ser Leu Val Trp Arg Pro
35 40 45
Asp Phe Ala Tyr Ser Asp Ser Arg Val Gly Phe Val Arg Gly Tyr
50 55 60
Ser Arg Arg Phe Trp Gln Gly Asp Thr Phe His Arg Gly Ser Asp
65 70 75
Lys Met Pro Gly Arg Val Val Thr Leu Leu Glu Asp His Glu Gly
80 85 90
Cys Thr Trp Gly Val Ala Tyr Gln Val Gln Gly Glu Gln Val Ser
95 100 105
Lys Ala Leu Lys Tyr Leu Asn Val Arg Glu Ala Val Leu Gly Gly
110 115 120
Tyr Asp Thr Lys Glu Val Thr Phe Tyr Pro Gln Asp Ala Pro Asp
125 130 135
Gln Pro Leu Lys Ala Leu Ala Tyr Val Ala Thr Pro Gln Asn Pro
140 145 150
Gly Tyr Leu Gly Pro Ala Pro Glu Glu Ala Ile Ala Thr Gln Ile
155 160 165
Leu Ala Cys Arg Gly Phe Ser Gly His Asn Leu Glu Tyr Leu Leu
170 175 180
Arg Leu Ala Asp Phe Met Gln Leu Cys Gly Pro Gln Ala Gln Asp
185 190 195
Glu His Leu Ala Ala Ile Val Asp Ala Val Gly Thr Met Leu Pro
200 205 210
Cys Phe Cys Pro Thr Glu Gln Ala Leu Ala Leu Val
215 220

<210> 17
<211> 111
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5378485CD1

<400> 17
Met Leu Ser Ala Leu Pro Gly Trp Gly Pro Ala His Leu Gln Arg

```
  1               5                  10                  15
Pro Leu Leu Gly Pro Ala Ser Cys Leu Gly Ile Leu Arg Pro Ala
                20                  25                  30
Met Thr Ala His Ser Phe Ala Leu Pro Val Ile Ile Phe Thr Thr
                35                  40                  45
Phe Trp Gly Leu Val Gly Ile Ala Gly Pro Trp Phe Val Pro Lys
                50                  55                  60
Gly Pro Asn Arg Gly Val Ile Ile Thr Met Leu Val Ala Thr Ala
                65                  70                  75
Val Cys Cys Tyr Leu Phe Trp Leu Ile Ala Ile Leu Ala Gln Leu
                80                  85                  90
Asn Pro Leu Phe Gly Pro Gln Leu Lys Asn Glu Thr Ile Trp Tyr
                95                 100                 105
Val Arg Phe Leu Trp Glu
               110
```

<210> 18
<211> 1303
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 961344CB1

<400> 18

```
cccacgcgtc cgcccacgcg tccgcaagat atggatgcta acaggcgata aactcgagac  60
agctacctgc attgccaaaa gttcacatct cgtgtctaga acacaagata ttcatatttt 120
cagacaggta accagtcggg gagaggcaca tttggagctg aatgcatttc gaaggaagca 180
tgattgtgca ctagtcatat ctggggactc tctggaggtt tgtctaaagt actacgagca 240
tgaatttgtg gagctggcct gccagtgccc tgccgtggtt tgctgccgct gctcacccac 300
ccagaaggcc cgcattgtga cactgctgca gcagcacaca gggagacgca cctgcgccat 360
cggtgatgga ggaaatgatg tcagcatgat tcaggcagca gactgtggga ttgggattga 420
gggaaaggag ggtaaacagg cctcgctggc ggccgacttc tccatcacgc agttccggca 480
cataggcagg ctgctcatgg tgcacgggcg gaacagctac aagaggtcgg cggcactcgg 540
ccagttcgtc atgcacaggg gccttatcat ctccaccatg caggctgtgt tttcctcagt 600
cttctacttc gcatccgtcc ctttgtatca gggcttcctc atggtggggt atgccaccat 660
atacaccatg ttcccagtgt tctccttagt gctggaccag gacgtgaagc cagagatggc 720
gatgctctac ccggagctgt acaaggacct caccaaggga agatccttgt ccttcaaaac 780
cttcctcatc tgggttttaa taagtattta ccaaggcggc atcctcatgt atggggccct 840
ggtgctcttc gagtctgagt tcgtccacgt ggtggccatc tccttcaccg cactgatcct 900
gaccgagctg ctgatggtgg cgctgaccgt ccgcacgtgg cactggctga tggtggtggc 960
cgagttcctc agcttaggct gctacgtgtc ctcactcgct tttctcaatg aatattttgg 1020
tataggcaga gtgtcttttg gagctttctt agatgttgcc tttatcacca ccgtgacctt 1080
cctgtggaaa gtgtcggcga tcaccgtggt cagctgcctc ccgctgtatg tcctcaagta 1140
cctgaggcgc aagctctctc ctcccagcta ctgcaagctg gcctcctaag gggctgtgca 1200
cccccagcgg gctggcccca gcaccttctg cccttcccag caccttgtgc ccttgccagt 1260
gaacgcaggg tttgccattg ctaccaagca agcaccacaa gaa                   1303
```

<210> 19
<211> 3395
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3128782CB1

<400> 19
cggaaatagc accgggcgcc gccacagtag ctgtaactgc caccgcgatg ccgaaggcgc 60
ccaagcagca gccgccggag cccgagtgga tcggggacgg agagagcacg agcccatcag 120
acaaagtggt gaagaaaggg aagaaggaca agaagatcaa aaaaacgttc tttgaagagc 180
tggcagtaga agataaacag gctggggaag aagagaaagt gctcaaggag aaggagcagc 240
agcagcagca acagcaacag cagcagcaaa aaaaaaagcg agatacccga aaaggcaggc 300
ggaagaagga tgtggatgat gatggagaag agaaagagct catggagcgt cttaagaagc 360
tctcagtgcc aaccagtgat gaggaggatg aagtacccgc cccaaaaccc cgcggaggga 420
agaaaaccaa gggtggtaat gtttttgcag ccctgattca ggatcagagt gaggaagagg 480
aggaggaaga aaaacatcct cctaagcctg ccaagccgga gaagaatcgg atcaataagg 540
ccgtatctga ggaacagcag cctgcactca agggcaaaaa gggaaaggaa gagaagtcaa 600
aagggaaggc taagcctcaa aataaattcg ctgctctgga caatgaagag gaggataaag 660
aagaagaaat tataaaggaa aaggagcctc ccaaacaagg gaaggagaag gccaagaagg 720
cagagcaggg ttcagaggaa gaaggagaag gggaagaaga ggaggaggaa ggaggagagt 780
ctaaggcaga tgatccctat gctcatctta gcaaaaagga gaagaaaaag ctgaaaaaac 840
agatggagta tgagcgccaa gtggcttcat taaaagcagc caatgcagct gaaaatgact 900
tctccgtgtc ccaggcggag atgtcctccc gccaagccat gttagaaaat gcatctgaca 960
tcaagctgga gaagttcagc atctccgctc atggcaagga gctgttcgtc aatgcagacc 1020
tgtacattgt agccggccgc cgctacgggc tggtaggacc caatggcaag ggcaagacca 1080
cactcctcaa gcacattgcc aaccgagccc tgagcatccc tcccaacatt gatgtgttgc 1140
tgtgtgagca ggaggtggta gcagatgaga caccagcagt ccaggctgtt cttcgagctg 1200
acaccaagcg attgaagctg ctggaagagg agcggcggct tcagggacag ctggaacaag 1260
gggatgacac agctgctgag aggctagaga aggtgtatga ggaattgcgg gccactgggg 1320
cggcagctgc agaggccaaa gcacggcgga tcctggctgg cctgggcttt gaccctgaaa 1380
tgcagaatcg acccacacag aagttctcag ggggctggcg catgcgtgtc tccctggcca 1440
gggcactgtt catggagccc acactgctga tgctggatga gcccaccaac cacctggacc 1500
tcaacgctgt catcttggctt aataactacc tccaggctg gcggaagacc ttgctgatcg 1560
tctcccatga ccagggcttc ttggatgatg tctgcactga tatcatccac ctcgatgccc 1620
agcggctcca ctactatagg ggcaattaca tgaccttcaa aaagatgtac cagcagaagc 1680
agaaagaact gctgaaacag tatgagaagc aagagaaaaa gctgaaggag ctgaaggcag 1740
gcgggaagtc caccaagcag gcggaaaaac aaacgaagga agccctgact cggaagcagc 1800
agaaatgccg acggaaaaac caagatgagg aatcccagga ggcccctgag ctcctgaagc 1860
gccctaagga gtacactgtg cgcttcactt ttccagaccc cccaccactc agccctccag 1920
tgctgggtct gcatggtgtg acattcggct accagggaca gaaaccactc tttaagaact 1980
tggattttgg catcgacatg gattcaagga tttgcattgt gggccctaat ggtgtgggga 2040
agagtacgct actcctgctg ctgactggca agctgacacc gacccatggg gaaatgagaa 2100
agaaccaccg gctgaaaatt ggcttcttca accagcagta tgcagagcag ctgcgcatgg 2160
aggagacgcc cactgagtac ctgcagcggg gcttcaacct gccctaccag gatgcccgca 2220
agtgcctggg ccgcttcggc ctggagagtc acgcccacac catccagatc tgcaaactct 2280
ctggtggtca gaaggcgcga gttgtgtttg ctgagctggc ctgtcgggaa cctgatgtcc 2340
tcatcttgga cgagccaacc aataacctgg acatagagtc tattgatgct ctaggggagg 2400
ccatcaatga atacaagggt gctgtgatcg ttgtcagcca tgatgcccga ctcatcacag 2460
aaaccaattg ccagctgtgg gtggtggagg agcagagtgt tagccaaatc gatggtgact 2520
ttgaagacta caagcgggag gtgttggagg ccctgggtga agtcatggtc agccggcccc 2580
gagagtgaag ctttccttcc cagaagtctc ccgagagaca tatttgtgtg gcctagaagt 2640
cctctgtggt ctccctctct ctgaagactg cctctggcct gcagctgacc tggcaaccat 2700
tcaggcacat gaaggtggag tgtgaccttg atgtgaccgg gatcccactc tgattgcatc 2760
catttcttctg aaagacttgt ttgttctgct tctcttcata taactgagct ggccttatcc 2820
ttggcatccc cctaaacaaa caagaggtga ccacctttatt gtgaggttcc atccagccaa 2880
gtttatgtgg cctattgtct caggactctc atcactcaga agcctgcctc tgatttaccc 2940
tacagcttca ggcccagctg ccccccagtc tttgggtggt gctgttcttt tctggtggat 3000

ttaatgctga ctcactggta caaacagctg ttgaagctca gagctggagg tgagctttctg 3060
aggcctttgc cattatccag cccaagattt ggtgcctgca gcctcttgtc tggttgagga 3120
cttggggcag gaaaggaatg ctgctgaact tgaatttccc tttacaaggg gaagaaataa 3180
aggaaaggag ttgctgccga cctgtcactg tttggagatt gatgggagtt ggaactgttc 3240
tcagtcttga tttgctttat tcagtttttct agcagctttt aatagtcccc tcttccccac 3300
taaatggatc ttgtttacag tattactgac agtgtttact gtttaaggat cataggattc 3360
cttaacccca accattcccg caaggaataa gcaat 3395

<210> 20
<211> 2549
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1720440CB1

<400> 20
acacgtccgc atgcaagacc atcaggcgcg atatctttgg cgcctggtcc tcctgaaggt 60
gaaggcaaag gtgcgacagt gcctgcagga gcggcggaca gtgcccattt tgtttgcctc 120
taccgttcgg cgccaccccg acaagacggc cctgatcttc gagggcacag atacccactg 180
gaccttccgc cagctggatg agtactcaag cagtgtagcc aacttcctgc aggcccgggg 240
ctgaccatcg gcgatgtggc tgccatcttc atggagaacc gcaatgagtt cgtgggccta 300
tggctgggca tggccaagct cggtgtggag gcagccctca tcaacaccaa cctgcggcgg 360
gatgctctgc tccactgcct caccacctcg cgcgcacggg cccttgtctt tggcagcgaa 420
atggcctcag ccatctgtga ggtccatgcc agcctggacc cctcgctcag cctcttctgc 480
tctggctcct gggagcccgg tgcggtgcct ccaagcacag aacacctgga ccctctgctg 540
aaagatgctc ccaagcacct tcccagttgc cctgacaagg gcttcacaga taaactgttc 600
tacatctaca catccggcac cacagggctg cccaaggccg ccatcgtggt gcacagcagg 660
tattaccgca tggctgccct ggtgtactat ggattccgca tgcggcccaa cgacatcgtc 720
tatgactgcc tccccctcta ccactcagca ggaaacatcg tgggaatcgg ccagtgcctg 780
ctgcatggca tgacggtggt gattcggaag aagttctcag cctcccggtt ctgggacgat 840
tgtatcaagt acaactgcac gattgtgcag tacattggtg aactgtgccg ctacctcctg 900
aaccagccac cgcgggaggc agaaaaccag caccaggttc gcatggcact aggcaatggc 960
ctccggcagt ccatctggac caacttttcc agccgcttcc acatacccca ggtggctgag 1020
ttctacgggg ccacagagtg caactgtagc ctgggcaact tcgacagcca ggtgggggcc 1080
tgtggtttca atagccgcat cctgtcctcc gtgtacccca tccggttggt acgtgtcaac 1140
gaggacacca tggagctgat ccgggggccc gacggcgtct gcattccctg ccagccaggt 1200
gagccgggcc agctggtggg ccgcatcatc cagaaagacc ccctgcgccg cttcgatggc 1260
tacctcaacc agggcgccaa caacaagaag attgccaagg atgtcttcaa gaaggggcac 1320
caggcctacc ttactggtga tgtgctggtg atggacgagc tgggctacct gtacttccga 1380
gaccgcactg gggacacgtt ccgctggaaa ggtgagaacg tgtccaccac cgaggtggaa 1440
ggcacactca gccgcctgct ggacatggct gacgtggccg tgtatggtgt cgaggtgcca 1500
ggaaccgagg gccgggccgg aatggctgct gtggccagcc ccactggcaa ctgtgacctg 1560
gagcgctttg ctcaggtctt ggagaaggaa ctgcccctgt atgcgcgccc catcttcctg 1620
cgcctcctgc ctgagctgca caaaacagga acctacaagt tccagaagac agagctacgg 1680
aaggagggct ttgaccccgc tattgtgaaa gacccgctgt tctatctaga tgccacgaag 1740
ggccgctacg tcccgctgga ccaagaggcc tacagccgca tccaggcagg cgaggagaag 1800
ctgtgattcc ccccatccct ctgagggccg gcggatgctg gatccggagc cccaggttcc 1860
gccccagagc ggtcctggac aaggccagac caaagcaagc agggcctggc acctccatcc 1920
tgaggtgctg cccctccatc caaaactgcc aagtgactca ttgccttccc aacccttcca 1980
gaggctttct gtgaaagtct catgtccaag ttccgtcttc tgggctgggc aggccctctg 2040
gttcccaggc tgagactgac gggtttttctc aggatgatgt cttgggtgag ggtaggagga 2100
ggacaagggg tcaccgagcc cttcccagag agcagggagc ttataaatgg aaccagagca 2160

```
gaagtcccca gactcaggaa gtcaacagag tgggcaggga cagtggtagc atccatctgg 2220
tggccaaaga gaatcgtagc cccagagctg cccaagttca ctgggctcca ccccacctc 2280
caggaggga ggagaggacc tgacatctgt aggtggcccc tgatgcccca tctacagcag 2340
gaggtcagga ccacgcccct ggcctctccc cactcccca tcctcctccc tgggtggctg 2400
cctgattatc cctcaggcag ggcctctcag tccttgtggg tctgtgtcac ctccatctca 2460
gtcttggcct ggctatgagg ggaggaggaa tgggagaggg ggctcagggg ccaataaact 2520
ctgccttgag tcctcctaaa aaaaaaaaa                                  2549
```

<210> 21
<211> 2562
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2274290CB1

<400> 21

```
gcgggagcca acatagagcc ctcagtggga tgagggtgaa actgctattg ccggcggctc  60
ctgttttacc gcgtcagcat gctggtgcat ttatttcggg tcgggattcg gggtggccca 120
ttcccagcca ggctgctacc gcccctccgc ttccagacat tctcagctgt caggtactct 180
gatggctacc gcagctcctc cctcctccgg gccgtggccc acctgcggtc ccagctctgg 240
gcccacctcc ctcgagcccc cctagctccc agatggagcc cctctgcctg gtgctgggtt 300
gggggagccc tgctaggccc catggtactg agtaagcatc cccacctctg ccttgtggcc 360
ctgtgtgagg cagaagaggc ccctcctgcc agctccacac cccatgtcgt ggggtctcgc 420
tttaactgga agctctttctg gcagtttctg caccccccacc tgctggtcct gggggtagcc 480
gtcgtgctgg ccttgggtgc ggcactcgtg aatgtacaga tccccctgct cctgggccag 540
ctggtagagg tcgtggccaa gtacacaagg gaccacgtag ggagtttcat gactgagtcc 600
cagaatctca gcacccacct gcttatcctc tatggtgtcc agggactgct gaccttcggg 660
tacctggtgc tgctgtccca cgttggcgag cgcatggctg tggacatgcg gagggccctc 720
ttcagctccc tgctccgaca agacatcacc ttctttgacg ccaataagac agggcagctg 780
gtgagccgct tgacaactga cgtgcaggag tttaagtcat ccttcaagct tgtcatctcc 840
cagggggctgc gaagctgcac ccaggtggca ggctgcctgg tgtccctgtc catgctgtcg 900
acacgcctca cgctgctgct gatggtggcc acaccagccc tgatgggagt gggcaccctg 960
atgggctcag gcctccgaaa attgtctcgc cagtgtcagg agcagatcgc cagggcaatg 1020
ggcgtagcag acgaggccct gggcaatgtg cggactgtgc gtgccttcgc catggagcaa 1080
cgggaagagg agcgctatgg ggcagagctg gaagcctgcc gctgccgggc agaggagctg 1140
ggccgcggca tcgccttgtt ccaagggctt tccaacatcg ccttcaactg catggtcttg 1200
ggtaccctat ttattggggg ctcccttgtg gccggacagc agctgacagg gggagacctc 1260
atgtccttcc tggtggcctc ccagacagtg caaaggtcca tggccaacct ctctgtcctg 1320
tttgggcagg tggtccgggg gctgagtgca ggtgcccggg tctttgagta catggccctg 1380
aaccccgtca tcccactgtc tgggggctgc tgcgtcccca aagagcagct gcgtggctcc 1440
gttacatttc agaacgtctg cttcagctac ccctgccgcc ccggcttcga ggtgctgaaa 1500
gacttcaccc tgacgctgcc ccctggcaag atcgtggccc tcgtgggcca gtctggcgga 1560
ggaaagacca ccgtggcttc cctgctggag cgcttctacg accccacggc aggcgtggtg 1620
atgctggatg ggcgggacct gcgcaccctt gacccctcct ggctccgggg ccaggttgtc 1680
ggcttcatca gccaggagcc cgtcctgttt gggacgacca tcatggaaaa catccgcttt 1740
gggaagctgg aagcttccga tgaagaggtg tacacagccg cccgggaagc gaatgctcac 1800
gagttcatca ccagcttccc cgagggctac aacacggtcg tcggtgaacg gggcactacc 1860
ctgtctgggg gccagaagca gcgcctggcc atcgcccgag cccttatcaa gcagcccacg 1920
gtgctgatac tggatgaagc taccagcgcg ctggatgcag agtccgagcg ggttgtacag 1980
gaggccctgg accgggccag tgcaggccgc acggtgctgg taattgccca ccggctcagc 2040
actgtccgtg gggcccactg cattgtcgtc atggccgatg gccgtgtctg ggaggctggg 2100
acacatgaag agctcctgaa gaaaggcggg ctatacgccg agctcatccg gaggcaggcc 2160
```

ctggatgccc cgaggacagc ggccccccg cccaaaaagc cagaaggccc caggagccac 2220
cagcacaagt cctgagaagg gcccccgag gtgtggtcgc tgccaagcat cagtgttagg 2280
gctggggctc agcctggggg agcctactgg ggactgagcc cccaggaggg ccagcatgtg 2340
gagagtcgct gcggctgctc ctgctcacaa taaagccggg gccgagcagc tggcagggga 2400
ggccaatccc tccctcccct ccccagtcct gccggctgcc tccctcccac cagagtctgc 2460
cagagtcatt gggctgcaat gggcagagac agagttccac gagacacctc cactctattc 2520
tccctttgcc cagacccctc cagacctctc aagagacgtt ct 2562

<210> 22
<211> 2314
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2740029CB1

<400> 22
cgctggtctt catgcgccct agccctcttt cggggatact ggccgacccc ctcttccttt 60
tcccctttag tgaaggcctc ccccgtcgcc gcgcggcttc ccggagccga ctgcagactc 120
cctcagcccg gtgttccccg cgtccggacg ccgaggtcgc ggcttcgcag aaactcgggc 180
ccctccatcc gccctcagaa aagggagcga tgttgatctc aggaagcaca aagggacctt 240
cctagctctg actgaaccac ggagctcacc ttggacagta tcactccgtg gaggaagact 300
gtgagactgt ggctggaagc cagattgtag ccacacatcc gccctgccc taccccagag 360
ccctggagca gcaactggct gcagatcaca gacacagtga ggatatgagt gtaggggtga 420
gcacctcagc ccctctttcc ccaacctcgg gcacaagcgt gggcatgtct accttctcca 480
tcatggacta tgtggtgttc gtcctgctgc tggttctctc tcttgccatt gggctctacc 540
atgcttgtcg tggctggggc cggcatactg ttggtgagct gctgatggcg gaccgcaaaa 600
tgggctgcct tccggtggca ctgtccctgc tggccacctt ccagtcagcc gtggccatcc 660
tgggtgtgcc gtcagagatc taccgatttg ggacccaata ttggttcctg ggctgctgct 720
actttctggg gctgctgata cctgcacaca tcttcatccc cgttttctac cgcctgcatc 780
tcaccagtgc ctatgagtac ctggagcttc gattcaataa aactgtgcga gtgtgtggaa 840
ctgtgacctt catctttcag atggtgatct acatgggagt tgtgctctat gctccgtcat 900
tggctctcaa tgcagtgact ggcttttgatc tgtggctgtc cgtgctggcc ctgggcattg 960
tctgtaccgt ctatacagct ctgggtgggc tgaaggccgt catctggaca gatgtgttcc 1020
agacactggt catgttcctc gggcagctgg cagttatcat cgtggggtca gccaaggtgg 1080
gcggcttggg gcgtgtgtgg gccgtggctt cccagcacgg ccgcatctct gggtttgagc 1140
tggatccaga ccccttttgtg cggcacacct tctggacctt ggccttcggg ggtgtcttca 1200
tgatgctctc cttatacggg gtgaaccagg ctcaggtgca gcggtacctc agttcccgca 1260
cggagaaggc tgctgtgctc tcctgttatg cagtgttccc cttccagcag gtgtccctct 1320
gcgtgggctg cctcattggc ctggtcatgt tcgcgtatta ccaggagtat cccatgagca 1380
ttcagcaggc tcaggcagcc ccagaccagt tcgtcctgta ctttgtgatg gatctcctga 1440
agggcctgcc aggcctgcca gggctcttca ttgcctgcct cttcagcggc tctctcagca 1500
ctatatcctc tgcttttaat tcattggcaa ctgttacgat ggaagacctg attcgacctt 1560
ggttccctga gttctctgaa gcccgggcca tcatgctttc cagaggcctt gcctttggct 1620
atgggctgct ttgtctagga atggcctata tttcctccca gatgggacct gtgctgcagg 1680
cagcaatcag catctttggc atggttgggg gaccgctgct gggactcttc tgccttggaa 1740
tgttctttcc atgtgctaac cctcctggtg ctgttgtggg cctgttggct gggctcgtca 1800
tggccttctg gattggcatc gggagcatcg tgaccagcat gggctccagc atgccaccct 1860
ctccctctaa tgggtccagc ttctccctgc ccaccaatct aaccgttgcc actgtgacca 1920
cactgatgcc cttgactacc ttctccaagc ccacagggct gcagcggttc tattccttgt 1980
cttactatag gtacagtgct cacaactcca ccacagtgat tgtggtgggc ctgattgtca 2040
gtctactcac tgggagaatg cgaggccggt ccctgaaccc tgcaaccatt tacccagtgt 2100
tgccaaagct cctgtccctc cttccgttgt cctgtcagaa gcggctccac tgcaggagct 2160

acggccagga ccacctcgac actggcctgt ttcctgagaa gccgaggaat ggtgtgctgg 2220
gggacagcag agacaaggag gccatggccc tggatggcac agcctatcag gggagcagct 2280
ccacctgcat cctccaggag acctccctgt gatg 2314

<210> 23
<211> 2155
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2414415CB1

<400> 23
gtggttccta tttcggaaaa ggacgttcta attcaaagct ctctcccaat atatttacac 60
gaatacgcat ttagaaaggg aggcagcttt tgaggttgca atcctactga gaaggatgga 120
agaaggagcc aggcaccgaa acaacaccga aaagaaacac ccaggtgggg gcgagtcgga 180
cgccagcccc gaggctggtt ccggaggggg cggagtagcc ctgaagaaag agatcggatt 240
ggtcagtgcc tgtggtatca tcgtagggaa catcatcggc tctggaatct ttgtctcgcc 300
aaagggagtg ctggagaatg ctggttctgt gggccttgct ctcatcgtct ggattgtgac 360
gggcttcatc acagttgtgg gagccctctg ctatgctgaa ctcggggtca ccatccccaa 420
atctggaggt gactactcct atgtcaagga catcttcgga ggactggctg ggttcctgag 480
gctgtggatt gctgtgctgg tgatctacca caccaaccag gctgtcatcg ccctcacctt 540
ctccaactac gtgctgcagc cgctcttccc cacctgcttc cccccagagt ctggccttcg 600
gctcctggct gccatctgct tattgctcct cacatgggtc aactgttcca gtgtgcggtg 660
ggccacccgg gttcaagaca tcttcacagc tgggaagctc ctggccttgg ccctgattat 720
catcatgggg attgtacaga tatgcaaagg agagtacttc tggctggagc caaagaatgc 780
atttgagaat ttccaggaac ctgacatcgg cctcgtcgca ctggctttcc ttcagggctc 840
ctttgcctat ggaggctgga actttctgaa ttacgtgact gaggagcttg ttgatcccta 900
caagaacctt cccagagcca tcttcatctc catcccactg gtcacatttg tgtatgtctt 960
tgccaatgtc gcttatgtca ctgcaatgtc cccccaggag ctgctggcat ccaacgccgt 1020
cgctgtgact tttggagaga agctcctagg agtcatggcc tggatcatgc ccatttctgt 1080
tgccctgtcc acatttggag gagttaatgg gtctctcttc acctcctctc ggctgttctt 1140
cgctggagcc cgagagggcc accttcccag tgtgttggcc atgatccacg tgaagcgctg 1200
caccccaatc ccagccctgc tcttcacatg catctccacc ctgctgatgc tggtcaccag 1260
cgacatgtac acactcatca actacgtggg cttcatcaac tacctcttct atggggtcac 1320
ggttgctgga cagatagtcc ttcgctggaa gaagcctgat atccccgcgc ccaatcaagat 1380
caacctgctg ttccccatca tctacttgct gttctgggcc ttcctgctgg tcttcagcct 1440
gtggtcagag ccggtggtgt gtggcattgg cctggccatc atgctgacag gagtgcctgt 1500
ctatttcctg ggtgtttact ggcaacacaa gcccaagtgt ttcagtgact tcattgagct 1560
gctaaccctg gtgagccaga agatgtgtgt ggtcgtgtac cccgaggtgg agcggggctc 1620
agggacagag gaggctaatg aggacatgga ggagcagcag cagcccatgt accaacccac 1680
tcccacgaag gacaaggacg tggcgggcca gccccagccc tgaggaccac cattccctgg 1740
ctactctctc cttcctcccc cttttatcct acctccctgc cttggtcccg ccaacacatg 1800
cgagtacaca cacacccctc tctctgcttt tgtcaggcag tggtaggact ttggtgtggg 1860
tggtgagaaa ttgtaaacaa aaactgacat tcatacccaa agaaccagcc tctcacccca 1920
gggtccatgt cccaggcccc actccagtgc tgcccacact cccagctgct ggaggagagg 1980
ggagatgcca aggtgccctg caggacctcc ctccgggcca cacctcagc tgcctcttca 2040
ggaaccggag ctcattactg ccttccctcc cagggaggcc ccttcagaga ggagaggcca 2100
caggagctgc attgtggggg gacaggctca agcaattctg tccccatcaa ggggt 2155

<210> 24
<211> 1475

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2466714CB1

<400> 24
ggagcgcagg gcaggggtag aggctcgtag atggaactgg tagtcagctg gagagcagca 60
tggaggcgtc ctgggggagc ttcaacgctg agcggggctg gtatgtctct gtccagcagc 120
ctgaagaagc ggaggccgaa gagttgagtc cgttgctaag caacgaactt cacagacagc 180
gatccccagg tgtttcattt ggtttatcag tgtttaattt gatgaatgcc atcatgggaa 240
gtggcatcct tggcttagct tatgttatgg ctaataccgg tgtctttgga tttagcttct 300
tgctgctgac agttgctctc ctggcttctt actcagtcca tcttctgctt agtatgtgta 360
ttcagacagc tgtaacatct tatgaagatc ttggactctt tgcatttgga ttacctggaa 420
agttggtggt ggcaggcacc ataataattc agaatattgg agctatgtca tcttatcttt 480
taattattaa aacagagctt cctgctgcta ttgcagaatt tttgactgga gactataata 540
gatattggta tcttgatgga caaacactac taataatcat atgtgttggc attgtgttcc 600
ctcttgcact tcttcccaaa ataggctttc ttggctacac aagtagttta tcatttttct 660
ttatgatgtt ctttgctctt gtggtaataa ttaaaaaatg gtccatccct tgtcctctga 720
cattaaatta tgtagagaaa ggcttccaga tttcaaatgt tacagatgat tgtaagccaa 780
agctctttca tttctccaaa gagagtgctt atgccttacc aaccatggct ttttcatttc 840
tctgccatac ctcaatattg cccatatact gtgaacttca aagtccttca aagaaaagaa 900
tgcagaatgt taccaataca gcaattgctt taagttttct catttatttt atatctgcac 960
tctttgggta cctcactttt tatgacaaag tggagtcaga attactaaaa ggttatagta 1020
aatacttatc acatgatgtt gttgtcatga ctgtgaagtt atgcatacta tttgctgtgc 1080
ttttgacagt ccctctaatc cacttccctg ccagaaaagc tgtaacaatg atgtttttct 1140
ccaattttcc attctcatgg attcgccatt ttttgatcac tctagcactc aatattatca 1200
tcgttttact tgcaatatat gttcctgaca ttagaaatgt atttggtgta gttggtgcca 1260
gtacatcaac atgtttgatt tttatattcc caggactatt ttatcttaaa cttagcagag 1320
aggatttttt gtcatggaaa aagcttgggg cattcgtttt gctcatcttt ggaatttttg 1380
ttgggaattt tagtttagca ctcatcattt ttgattggat taataaataa aagaaatatt 1440
ttcctacttc ttacaagaat aataaaaaaa aaaaa 1475

<210> 25
<211> 1793
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2617942CB1

<400> 25
gcggtggcta cggcgacggg agccggcggc gctgcgggtc agcggtcgcg taggacccag 60
cggactcggc agcctggggc gcccggcgga gctgaaccgc ggcccccggt ggtgggctca 120
gccggtcgag ctgcgcggga ggcaaatgaa gataaaacaa tgttcgccaa cctaaaatac 180
gtttccctgg gaattttggt ctttcagact accagtttgg ttctaacaat gcgttattcc 240
agaactttaa aagaagaagg acctcgttat ctatcttcta cagcagtggt tgttgctgaa 300
cttttgaaga taatggcctg cattttattg gtctacaaag acagcaaatg tagtctaaga 360
gcactgaatc gagtactaca tgatgaaatt cttaataaac ctatggaaac acttaaactt 420
gctattccat cagggatcta tactcttcag aataatttac tgtatgtggc actaaaatac 480
ctagatgcag ctacttatca ggtcacgtat cagttaaaaa ttcttacaac agcattattt 540
tctgtgtcta tgcttagtaa aaaattgggt gtataccagt ggctgtccct agtaattttg 600

atgacaggag ttgcttttgt acagtggccc tcagattctc agcttgattc taaggaactt 660
tcagctggtt ctcaatttgt aggactcatg gcagttctca cagcatgttt ttcaagtggc 720
tttgctgggg tttactttga gaaaatctta aaagaaacaa aacaatcagt gtggataaga 780
aatattcagc ttggtttctt tggaagtata tttggattaa tgggtgtata catttatgat 840
ggagaactgg tatcaaagaa tggatttttt cagggatata accgactgac ctggatagta 900
gttgttcttc aggcacttgg aggccttgta atagctgctg ttattaagta tgcagataat 960
attttaaaag gatttgcaac ctctttatcg ataatattat caacattgat ctcctatttt 1020
tggcttcaag atttgtgcc aaccagtgtc tttttccttg gagccatcct tgtaataaca 1080
gctacttttt tgtatggtta tgatcccaaa cctgcaggaa atcccactaa agcatagttg 1140
tatactatct ttaactggtt tttcacgatg gggcactagg aatctcgaca ttaatcttgc 1200
acagaggact tctacagagt ctgagaagat atcatcatgc tgaatctgat catactgttt 1260
tttaaaagtt taaggataag acatgtgtat atgtaacaaa acacattgca tctagaaatc 1320
aaaacttgaa agtatttcca gggattagga ttagaaggaa tattagagga aacttgaaat 1380
ctgagtttaa aaagatttta ccttttttgat tgctgcagaa atgtcctatg cactctttgc 1440
aagagcacac aacaaatgtc agataccaat ttttgcaaat tagatttaat cttattaaat 1500
gtttttatct tactctttct gtacagatat atcaaatcac atgaaatatt taaagtttga 1560
aaattataat tacctataaa gctgtgaaaa atagaagtat aatttgaaaa aacatttcac 1620
ttatcagaga tttttatatt tatacaaaag attacttaat gaaggattgc taaatgtttt 1680
tgggtcaatt accttaagat taatattccg ggtctgatct gtcagggaat aaatatcaaa 1740
tctaaatttt aatgtggggg ttcatactat ttctcccata agaattttag ggt 1793

<210> 26
<211> 1141
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2945431CB1

<400> 26
ctggtttgtg tgtgtgcacc tccgtgaaat gtaggcacct tgaggacaga gtccagcctt 60
tggttcttt ggtattgctc atagcactgg cacagttcta ggtacccagc tactaacaga 120
tcatttggtg gggatggggt ggggagcaga gtggggttat gttcaggtct catacccagg 180
ctttcatgga ggtgctagcc ctgtagtcag aaactgagct gggagcagaa gtggctacat 240
ctccaaccac tagactccat gtcattgtcc cccagatccc agttggctat catcccccag 300
gagccctttt tgttcagtgg gactgttcgg gaaaacctgg acccccaggg cctacataag 360
gacagggcct tgtggcaggc cctgaagcag tgccacctga gtgaggtgat tacatccatg 420
ggtggtctgg atggtgagct gggtgagggg ggccggagct tatctcttgg gcagaggcag 480
ctgttgtgtt tggccagggc tctcctcaca gatgccaaga tcctgtgtat cgatgaggcc 540
acagcaagtg tggaccagaa gacagaccag ctgctccagc agaccatctg caaacgcttt 600
gccaacaaga cagtgctgac cattgcccat aggctcaaca cgatcctgaa ctcagaccgg 660
gtgctggtgc tacaagcggg gagagtggta gagctggact ccccggccac cctgcgcaac 720
cagccccact ccctgttcca gcagctgctg cagagcagcc agcagggagt ccctgcctca 780
ctcggaggtc cctgagccca atcccacacc ctgcagagtt ctcccctctc tctgatccag 840
gccgggccta tacagaggtg ctggctgctt gtttacattc tcctctgggg ctctacctct 900
ccacacttcc ccagaaggga aaagggcacc ctggattact ctttggaaat cactccttgg 960
tgggcagcat cctgaggctt ccccagaacc aggcctctgc tctggccctc ttgcatctgg 1020
aacgccaggt gggtttttct ggcataggag cccacttgca ttttcatagt tttatttgat 1080
aaaattccat cttacattct gtgtattaaa aaaataatat ttctggtgtg agaaaaaaaa 1140
a 1141

<210> 27

<211> 1371
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4074113CB1

<400> 27
gactggatgg attgatgggt ggatatatag aaaggtagac agatggaaag agagatggaa 60
ggtagacctt tacacaatga gggatggata gacagatctc gggtacagca gaaagatctc 120
cccaataaat gcccacaaac cctctggtca gagcaggcct ttcctcccaa ccccgggcag 180
gtgggcatcg ttggcaggac cggggcaggg aagtcctccc tggccagtgg gctgctgcgg 240
ctcccagagg cagctgaggg tgggatctgg atcgacgggg tccccattgc ccacgtgggg 300
ctgcacacac tgcgctccag gatcagcatc atcccccagg acccccatcct gttcccctggc 360
tctctgcgga tgaacctcga cctgctgcag gagcactcgg acgaggctat ctgggcagcc 420
ctggagacgg tgcagctcaa agccttggtg gccagcctgc ccggccagct gcagtacaag 480
tgtgctgacc gaggcgagga cctgagcgtg ggccagaaac agctcctgtg tctggcacgt 540
gcccttctcc ggaagaccca gatcctcatc ctggacgagg ctactgctgc cgtggaccct 600
ggcacggagc tgcagatgca ggccatgctc gggagctggt ttgcacagtg cactgtgctg 660
ctcattgccc accgcctgcg ctccgtgatg gactgtgccc gggttctggt catggacaag 720
gggcaggtgg cagagagcgg cagcccggcc cagctgctgg cccagaaggg cctgttttac 780
agactggccc aggagtcagg cctggtctga gccaggaccc tcaaccgtac cccagttgga 840
ccagcccgca cagcctgcag tgctggagat ggaagtgacc cgtggtcatc gatagctcca 900
cacgatattg agtctagacc tgtgtttgct ctctgggagg aaaatggcag agaaagtggc 960
caattatcac agagcatcag agccggaagg acctagcaat acacaggtct gcccgggcag 1020
ggccatctc gccctgtcca ccctgcagcc aatgtcaaca gcgactctca gccccgctgt 1080
actctggact cacctggggg cctcaagcac atgcccaggc tcccggctag acccttaaat 1140
cagaatctct gaggctggga actgccatgc tgtgtgtact ttttacaaat taacactttt 1200
attttgggat aatcccagac tcacatgcag ttaaagaaac aataatatag agagattcgt 1260
gtacttggta ccccatttca cccaatggta acatcttgca aaactctagg ataaagcatc 1320
acagccaggg tgttgacatt gacacaacaa tcttgctcgg atgtccgcga g 1371

<210> 28
<211> 2752
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1413743CB1

<400> 28
gggaaagaat cccccaagct ccatttcatg agtaagcgtg agagccgctc agtttcctcc 60
agctctgctg aagccagcac agaagtagcc caaactcttc cctctgctga cagcaaattt 120
taggcaaagt cttgagaaag aagaaattgg gtccagaaag ggaagtgagg agaatcagat 180
cccagacctt tggggagaag gagcaaccgc ctctggcaca gcccatcagg gagaaagagc 240
aggttgagaa gagtcctaag ctaacagccc caaacaggtg ggtgttgctc agctccctga 300
ggcatgtggt tgtaaggcag aacccacaga ccttgcagga agaaggctct cggggccatg 360
gcccaggtca gcatcaacaa tgactacagc gagtgggact tgagcacgga tgccgggagg 420
cgggctcggc tgctgcagag tccctgtgtg gacacagccc ccaagagtga gtgggaagcc 480
tctcctgggg gtctggacag aggcaccact tccacacttg gggccatctt catcgtcgtc 540
aacgcgtgcc tgggtgcagg gttactcaac ttcccagcag ccttcagcac tgcgggggcc 600
gtggcagcag gcatcgcact gcagatgggt atgctggttt tcatcatcag tggccttgtc 660

atcctggcct actgctccca ggccagcaat gagaggacct accaggaggt ggtatgggct 720
gtgtgtggca agctgacagg tgtgctatgt gaggtggcca tcgctgtcta caccttttggc 780
acctgcattg ccttcctaat catcattggc gaccagcagg acaagattat agctgtgatg 840
gcgaaagagc cggagggggc cagcggccct tggtacacag accgcaagtt caccatcagc 900
ctcactgcct tcctcttcat cctgcccctc tccatcccca gggagattgg tttccagaaa 960
tatgccagct tcctgagcgt cgtgggtacc tggtacgtca cagccatcgt tatcatcaag 1020
tacatctggc cagataaaga gatgacccca gggaacatcc tgaccaggcc ggcttcctgg 1080
atggctgttgt tcaatgccat gcccaccatc tgcttcggat ttcagtgcca cgtcagcagt 1140
gtgcccgtct tcaacagcat gcagcagcct gaagtgaaga cctggggtgg agtggtgaca 1200
gctgccatgg tcatagccct cgctgtctac atgggtacag gcatctgtgg cttcctgacc 1260
tttggagctg ctgtggatcc tgacgtgctc ctgtcctatc cctcggagga catggccgtg 1320
gccgttgccc gagccttcat catcctgagc gtgctcacct cctaccctat cctgcacttc 1380
tgtgggcggg cggtggtgga aggcctgtgg ctgccgctacc agggggtgcc agtggaggag 1440
gacgtggggc ggagcggcg gcggcgagtg ctgcagacgc tggtctggtt cctgctcacc 1500
ctgctgctgg cgctcttcat ccctgacatc ggcaaggtga tctcagtcat tggaggcctg 1560
gccgcctgct tcatcttcgt cttcccaggg ctgtgcctca ttcaagccaa actctctgag 1620
atggaagagg tcaaaccagc cagctggtgg gtgctggtca gctacggagt cctcttggtc 1680
accctgggag ccttcatctt cggccagacc acagccaacg ccatctttgt ggatctcttg 1740
gcataaccac tgcctcccag ggaacacaag gcctttgcca ttggtcgcag gaacccatct 1800
cttagagcta tggggccatt cttagtccac gatcattcca actggtggga tgacatccgg 1860
acatcctctt ccagggactg gggcaaactc aggccccaca cctctggaca gctcaaatcc 1920
agtccccttc ctgctcccca gtcctggcag tgccgtggat ggcggcagga agtctcacat 1980
catagaggac ccctcctcct ctcccagttc tcaacttctc catgcctgga atccacgggt 2040
gaagagagtc ggtagatctc ataagaaaga atccagtctg acttccctct ggagaatgac 2100
tatggacaga aggccaccat cctccacaga gcaccctgtc ctgagtaggg gttgtgctca 2160
ttaccccagg ccagtggtag cttcctcagg agcctggcca cttccaacgg tagcactgaa 2220
gtcatgcaaa tgcatagtca ggtagattca gacctttgtcc acaccttcct gggcaacccc 2280
caccatgaac ctgtcagcct ctttcccata gctaatagac atttcccagg ccttgagggg 2340
ccccaccctg tctctttcat caaacctgat ggtccaggct gggcatccct ctcctcctcc 2400
atccccagac atcaccaggt ctaatgttta caaacggtgc cagcccggct ctgaagccaa 2460
gggccgtccc gtgccacggt gctgtgagta ttcctccgtt agctttcccc atgaaggttgg 2520
gagtatctgc ttttgtgtct gagatgggcc cctcttttca gaggccgcag ggtgggtgat 2580
ggagaaggct gagaaccttt cagaccctct gtgtgggctg ggctggtcag aatcagggtg 2640
tacctccccg acaccttctt tttcagtgat gttttctctt ctccctgcct ttcctctgcc 2700
tcctccctg ccagccctag cgtgactacc cagagacaaa aaaaaaaaaa aa 2752

<210> 29
<211> 2580
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1733477CB1

<400> 29
ggagcagccc gcaccggaca acttgcgagc catggggctg gcggatgcgt cgggaccgag 60
ggacacacag gcactgctgt ctgcaacaca agcaatggac ctgcggaggc gagactacca 120
catggaacgg ccgctgctga accaggagca tttggaggag ctggggcgct gggctcagc 180
acctaggacc caccagtggc ggacctggtt gcagtgctcc cgtgctcggg cctatgccct 240
tctgctccaa cacctcccgg ttttggtctg gttaccccgg tatcctgtgc gtgactggct 300
cctgggtgac ctgttatccg gcctgagtgt ggccatcatg cagcttccgc agggcttggc 360
ctacgccctc ctggctggat tgccccccgt gtttggcctc tatagctcct tctaccctgt 420
cttcatctac ttcctgtttg gcacttcccg gcacatctcc gtggggacct ttgctgtcat 480

gtctgtgatg gtgggcggtg tgacagaatc cctggccccg caggccttga acgactccat 540
gatcaatgag acagccagag atgctgcccg ggtacagggtg gcctccacac tcagtgtcct 600
ggttggcctc ttccaggtgg ggctgggcct gatccacttc ggcttcgtgg tcacctacct 660
gtcagaacct cttgtccgag gctataccac agctgcagct gtgcaggtct tcgtctcaca 720
gctcaagtat gtgttttggcc tccatctgag cagccactct gggccactgt ccctcatcta 780
tacagtgctg gaggtctgct ggaagctgcc ccagagcaag gttggcaccg tggtcactgc 840
agctgtggct ggggtggtgc tcgtggtggt gaagctgttg aatgacaagc tgcagcagca 900
gctgcccatg ccgataccgg gggagctgct cacgctcatc gggggccacag gcatctccta 960
tggcatgggt ctaaagcaca gatttgaggt agatgtcgtg ggcaacatcc ctgcagggct 1020
ggtgccccca gtggccccca acacccagct gttctcaaag ctcgtgggca gcgccttcac 1080
catcgctgtg gttgggtttg ccattgccat ctcactgggg aagatcttcg ccctgaggca 1140
cggctaccgg gtggacagca accaggagct ggtggccctg ggcctcagta accttatcgg 1200
aggcatcttc cagtgcttcc ccgtgagttg ctctatgtct cggagcctgg tacaggagag 1260
caccgggggc aactcgcagg ttgctggagc catctcttcc cttttcatcc tcctcatcat 1320
tgtcaaactt ggggaactct tccatgacct gcccaaggcg gtcctggcag ccatcatcat 1380
tgtgaacctg aagggcatgc tgaggcagct cagcgacatg cgctccctct ggaaggccaa 1440
tcgggcggat ctgcttatct ggctggtgac cttcacggcc accatcttgc tgaacctgga 1500
ccttggcttg gtggttgcgg tcatcttctc cctgctgctc gtggtggtcc ggacacagat 1560
gccccactac tctgtcctgg ggcaggtgcc agacacggat atttacagag atgtggcaga 1620
gtactcagag gccaaggaag tccgggggggt gaaggttcttc cgctcctcgg ccaccgtgta 1680
ctttgccaat gctgagttct acagtgatgc gctgaagcag aggtgtggtg tggatgtcga 1740
cttcctcatc tcccagaaga agaaactgct caagaagcag gagcagctga agctgaagca 1800
actgcagaaa gaggagaagc ttcggaaaca ggctgcctcc cccaagggcg cctcagtttc 1860
cattaatgtc aacaccagcc ttgaagacat gaggagcaac aacgttgagg actgcaagat 1920
gatggtgagc tcaggagata agatggaaga tgcaacagcc aatggtcaag aagactccaa 1980
ggccccagat gggtccacac tgaaggccct gggcctgcct cagccagact tccacagcct 2040
catcctggac ctgggtgccc tctcctttgt ggacactgtg tgcctcaaga gcctgaagaa 2100
tattttccat gacttccggg agattgaggt ggaggtgtac atggcggcct gccacagccc 2160
tgtggtcagc cagcttgagg ctgggcactt cttcgatgca tccatcacca agaagcatct 2220
ctttgcctct gtccatgatg ctgtcacctt tgccctccaa cacccgaggc ctgtccccga 2280
cagccctgtt tcggtcacca gactctgaac atgctacatc ctgcccaaga ctgcacctct 2340
ggaggtgcag ggcacccttg agaagcccct cacccctagg ccgcctccag gtgctaccca 2400
ggagtcccct ccatgtacac acacacaact cagggaagga ggtcctggga ctccaagttc 2460
agcgctccag gtctgggaca gggcctgcat gcagtcaggc tggcagtggc gcggtacagg 2520
gagggaactg gtgcatattt tagcctcagg aataaagatt tgtctgctca aaaaaaaaaa 2580

<210> 30
<211> 1481
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2641908CB1

<400> 30
tgatgattgt gctggtggtg gtgatcatga cagagacaac aataacaatc atcacatcgt 60
gatggtaatg tcgtgactaa atttgtcatt tagtcacaac gatatgggtg atgtgaatga 120
gggtgatatt taagctgaaa ggaatagaaa tgatgatgac agcaactcgc ccctctacct 180
cgggatcctg tttgcagtga ccatgatggg gccaggcctg gcctttgggc tgggcagcct 240
catgctgcgc ctttatgtgg acattaacca gatgccagaa ggtggtatca gcctgaccat 300
aaaggacccc cgatgggtgg gtgcctggtg gctgggtttc ctcatcgctg ccggtgcagt 360
ggccctggct gccatcccct acttcttctt ccccaaggaa atgcccaagg aaaaacgtga 420
gcttcagttt cggcgaaagg tcttagcagt cacagactca cctgccagga agggcaagga 480

ctctccctct aagcagagcc ctggggagtc cacgaagaag caggatggcc tagtccagat 540
tgcaccaaac ctgactgtga tccagttcat taaagtcttc cccagggtgc tgctgcagac 600
cctacgccac cccatcttcc tgctggtggt cctgtcccag gtatgcttgt catccatggc 660
tgcgggcatg gccaccttcc tgcccaagtt cctggagcgc cagttttcca tcacagcctc 720
ctacgccaac ctgctcatcg gctgcctctc cttcccttcg gtcatcgtgg gcatcgtggt 780
gggtggcgtc ctggtcaagc ggctccacct gggccctgtg ggatgcggtg ccctttgcct 840
gctggggatg ctgctgtgcc tcttcttcag cctgccgctc ttctttatcg gctgctccag 900
ccaccagatt gcgggcatca cacaccagac cagtgcccac cctgggctgg agctgtctcc 960
aagctgcatg gaggcctgct cctgcccatt ggacggcttt aaccctgtct gcgaccccag 1020
cactcgtgtg gaatacatca caccctgcca cgcaggctgc tcaagctggg tggtccagga 1080
tgctctggac aacagccaga gtcctcccac ctcccaccct catgctgggc atcagcatct 1140
aaacctgagg ctcctccagg gagagacctg ggctgcactg gctggtgcag aagaacctgt 1200
tgatggtgca tagtccttca gaagccagcc aggcacccac tgggcctgag agcccttcca 1260
gagaccccca ggccttggca ggtggagcag tgaactcctg tggatatggg aaccgattca 1320
aatccttctt aggcctctaa ctgactctgt taccttaggc aaattattta actagtgcct 1380
cagtttcttg gtctgtaaaa taggggagat attattaagt gcctactaca gagcaggaat 1440
gtgctgaata aatgctttac ctggatgaaa aaaaaaaaaa a 1481

<210> 31
<211> 667
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2656554CB1

<400> 31
ctaaagtggc agtgtttctt ctgaaatttct caggcagtca gactgtctta ggcaaatctt 60
gataaaatag cccttatcca ggtttttatc taaggaatcc caagaagact ggggaatgga 120
gagacagtca agggttatgt cagaaaagga tgagtatcag tttcaacatc agggagcggt 180
ggagctgctt gtcttcaatt ttttgctcat ccttaccatt ttgacaatct ggttatttaa 240
aaatcatcga ttccgcttct tgcatgaaac tggaggagca atggtgtatg acaagccgcc 300
gaaatttgcc atgtcacgag agcaaatgtc acagtcatgt tctcacacgg cacataatgc 360
aagtctgttg acagatgcgg gtccattgtc atgtggggag tcgagggcga gctgtttgtt 420
tttgtaatga tgttgggaag tgatggctct gcagtcacaa agagcagcct tctctcactg 480
gctgcaccga tgaacattac gaagttctag aaaaaacatc acttcaaaat gcctggagta 540
attcctctta tatcaactaa tttcaagaag aaaacttgca gaaactaacc ccacccctct 600
taagagaata ttgtgtccaa gtccttttta tttatacgaa cagtgtctta ttttcttata 660
atgaaat 667

<210> 32
<211> 1635
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2719228CB1

<400> 32
atagctgtct tgagccccaa gcctcttcct cccctgctgc ccctctgcag ccattcggga 60
tgggaccccc tctggggtgt cagcacgaaa gggctaacgg gagccccttc cttggctctc 120

```
ccctgtaggt tacagagcca tcacggtgcg gaagctcatg cagggcatgg gcctttggcct 180
ctccagcgtc tttgctctgt gcctgggcca cacctccagc ttctgtgagt ctgtggtctt 240
tgcatcagcc tccatcggcc tccagacctt caaccacagt ggcatttctg ttaacatcca 300
ggacttggcc ccgtcctgcg ccggctttct gtttggtgtg gccaacacag ccgggggcctt 360
ggcaggtgtc gtgggtgtgt gtctaggcgg ctacttgatg gagaccacgg gctcctggac 420
ttgcctgttc aaccttgtgg ccatcatcag caacctgggg ctgtgcacct tcctggtgtt 480
tggacaggct cagagggtgg acctgagctc tacccatgag gacctctagc tcccaacccc 540
acagcctctc caaggaccca ggcgccagca gccccgggac acaggggact cagtgtgtga 600
gacttggtca ctccatgtca gacacacgag cagagaggaa cacaaaccac tgtggagcct 660
gaagctcctt aagaagagtc cacaacagct ggtgggaggg tggggtgggc ctgggtccag 720
accaggctcg ctgctctctg ggcctcagtt tccccacctg ccagcgggct cggccctgtc 780
ctcctcacag gctggtgtgg ccgtcagggt gggtggggtt attgttagta ggcgcagcct 840
cattcccacc acgatctgtt ccgcgtggtt cccgccaaac ctccctcggt cgccgtgttc 900
tccgcaagcc tcctgcagcg cccgcctgcc aatgtgaggc tggcaccagg ctgcagcctc 960
cccaatccca gcccactttg ctgtgtctct ggcgggctgt cctccttggt gggagctgtc 1020
ctgcacactg taggatgctt aaaggtatcc ctggcctcca cccacccctа gccagcagct 1080
cccagtcaga caacagccag aaatgtctcc agactctgcc cagcctcccc aggtagccac 1140
cctcgagaca cgacctcaga gtctctgtgt ctcctagaag cctgacagag acccccaggg 1200
cagtgggtgg gtggcgggct agagaccctt gcctgtgtcc gggaccctgg cgccgctctc 1260
ccctcctgtg gatcctccg cactaacagt gttctcagtg ggcagacgcc tgggcacccc 1320
ttgggccctg cccagcatgg ccatggcgca ggctctcgaa cccgcatggc tttcccaggc 1380
ctggtgattc tgctctccag ggacggttgg caccttcctc gggggcgggc cccacgcacc 1440
ccagaacaca cagacccacc tttctggcgt tctttctacc tccctttcg ttgcctgagg 1500
agctggtggt ttcatgagtt aatgatacat cttgcaaggt gtacacatag agaaaaaaac 1560
ctaaaaatgt ggaaaagcac gccaaagcct tatttaaata ataactatta aactattcaa 1620
aaagaaaaaa aaaaa                                                  1635
```

<210> 33
<211> 1447
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3657824CB1

<400> 33

```
ccgagcggtg ccaggccagg tgtgtgcgtc cgtcggtctt tccgtgccca cgccggagac  60
cagccccgga ggccgcctgg gcctatccct gtgccaggca ccatgaagca ggagtctgca 120
gccccgaaca ccccgcccac ctcgcagtcc cctacgccgt ccgctcagtt ccccgaaaac 180
gacggcgacc ctcaagcgct gtggattttc gggtacggct ccctggtgtg gaggcccgac 240
ttcgcctaca gcgacagccg tgtgggcttc gtgcgcggct acagccgccg tttctggcag 300
ggagacacct tccatcgggg cagcgacaag atgcctggcc gtgtggtgac gctccttgaa 360
gatcatgagg gctgcacttg gggcgtggca taccaagtgc aaggggagca ggtaagcaag 420
gccctgaagt acctgaatgt gcgagaggca gtgcttggtg gctacgatac caaggaggtc 480
accttctatc cccaagatgc tcctgaccaa ccactgaagg cattggccta tgtggccacc 540
ccacagaacc ctggttacct gggccctgcg cctgaagagg ccattgccac gcgatcctg 600
gcctgccggg gcttctccgg ccacaacctt gaatacttgc tgcgtctggc agacttcatg 660
cagctctgtg ggcctcaggc gcaggacgag cacctggcag ccatcgtgga cgctgtgggc 720
accatgttgc cctgcttctg ccccaccgag caggctctgg cgctggtgtg aggggctgag 780
cccctgcggg gagtgctcat gtggacatca gggccagaca cccactccag tgcacaagac 840
agacttgcga ccgcttgagc ccactgagca gatatggtgg gtggctggag gcttctcttt 900
ctcagtccct gcctgtctgc cagcctgcag ctctcctgct tgacactgac ttactacttg 960
aaactttatt tattgcacca tgttggtgtg gtgggcaggt ggaggggcctg ccctggacac 1020
```

```
aggggccctg ctgagcagtg gccccatcct ggaacttgac cagattcccc ccagtgctgc 1080
tgctaacccc acaccaccca ggcctccacc tccccaggga gtctccaaga gcctcgatcc 1140
tctgctcact cagcccagcc atccatagcc ctgggaattc cacctgccaa ggatcccagc 1200
aggctggatg agggatagta gggcatgagg agaaggagcc ctgtaaggac tgaggccccg 1260
gccagcccct ctcctccacc agttccccag agcagagctg gagctgatgc ctggacacag 1320
ctgctgagcc tggcctgggc ctcttaccca cttggttgtt ttcttgtccc tctgtctgtc 1380
tgtctatcta cttgtctgtc tgggccactc ctgcctgtgt gttggtctat tcctgggaag 1440
ctcatca                                                          1447
```

<210> 34
<211> 657
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5378485CB1

<400> 34

```
gactcctgtt gcgcatgctc agcgcgctgc ccggctgggg acccgcgcac ctgcagcgcc  60
cgctgctcgg ccctgcatcc tgcctgggca tcctgcgccc ggccatgacg gcgcactcat 120
tcgccctccc ggtcatcatc ttcaccacgt tctggggcct cgtcggcatc gccgggccct 180
ggttcgtgcc gaagggaccc aaccgcggag tgatcatcac catgctggtc gccaccgccg 240
tctgctgtta cctctttctgg ctcatcgcca tcctggcgca gctgaacccc ctgttcgggc 300
cccagctgaa gaatgagacc atctggtacg tgcgcttcct gtgggagtga cccgccgccc 360
ccgacccagg tgcccagctc tcggaatgac tgtggctcca ctgtccctga caaccccttc 420
gtccggaccc tcccccacac aactatgtct ggtcaccagc tccctcctgc tggcacccag 480
agacccggac ccgcaggccc tgcctggttc ctggaagtct tcccagtctt cccagccagc 540
ccggggccct ggggagccct gggcacagca gcggccgagg ggatgtcctg ctccaatact 600
cgcactgctc tggagtttgc actctttcgc aaggagatgc tgctggggag ctggtat    657
```

<210> 35
<211> 646
<212> PRT
<213> Mus musculus

<300>
<308> GenBank ID No: g2612939

<400> 35

```
Met Arg Ala Pro Gly Ala Gly Thr Ala Ser Val Ala Ser Leu Ala
  1               5                  10                  15
Leu Leu Trp Phe Leu Gly Leu Pro Trp Thr Trp Ser Ala Ala Ala
             20                  25                  30
Ala Phe Cys Val Tyr Val Gly Gly Gly Gly Trp Arg Phe Leu Arg
         35                  40                  45
Ile Val Cys Lys Thr Ala Arg Arg Asp Leu Phe Gly Leu Ser Val
     50                  55                  60
Leu Ile Arg Val Arg Leu Glu Leu Arg Arg His Arg Arg Ala Gly
 65                  70                  75
Asp Thr Ile Pro Cys Ile Phe Gln Ala Val Ala Arg Arg Gln Pro
             80                  85                  90
Glu Arg Leu Ala Leu Val Asp Ala Ser Ser Gly Ile Cys Trp Thr
         95                 100                 105
```

```
Phe Ala Gln Leu Asp Thr Tyr Ser Asn Ala Val Ala Asn Leu Phe
                110                 115                 120
Arg Gln Leu Gly Phe Ala Pro Gly Asp Val Val Ala Val Phe Leu
                125                 130                 135
Glu Gly Arg Pro Glu Phe Val Gly Leu Trp Leu Gly Leu Ala Lys
                140                 145                 150
Ala Gly Val Val Ala Ala Leu Leu Asn Val Asn Leu Arg Arg Glu
                155                 160                 165
Pro Leu Ala Phe Cys Leu Gly Thr Ser Ala Ala Lys Ala Leu Ile
                170                 175                 180
Tyr Gly Gly Glu Met Ala Ala Ala Val Ala Glu Val Ser Glu Gln
                185                 190                 195
Leu Gly Lys Ser Leu Leu Lys Phe Cys Ser Gly Asp Leu Gly Pro
                200                 205                 210
Glu Ser Ile Leu Pro Asp Thr Gln Leu Leu Asp Pro Met Leu Ala
                215                 220                 225
Glu Ala Pro Thr Thr Pro Leu Ala Gln Ala Pro Gly Lys Gly Met
                230                 235                 240
Asp Asp Arg Leu Phe Tyr Ile Tyr Thr Ser Gly Thr Thr Gly Leu
                245                 250                 255
Pro Lys Ala Ala Ile Val Val His Ser Arg Tyr Tyr Arg Ile Ala
                260                 265                 270
Ala Phe Gly His His Ser Tyr Ser Met Arg Ala Ala Asp Val Leu
                275                 280                 285
Tyr Asp Cys Leu Pro Leu Tyr His Ser Ala Gly Asn Ile Met Gly
                290                 295                 300
Val Gly Gln Cys Val Ile Tyr Gly Leu Thr Val Val Leu Arg Lys
                305                 310                 315
Lys Phe Ser Ala Ser Arg Phe Trp Asp Asp Cys Val Lys Tyr Asn
                320                 325                 330
Cys Thr Val Val Gln Tyr Ile Gly Glu Ile Cys Arg Tyr Leu Leu
                335                 340                 345
Arg Gln Pro Val Arg Asp Val Glu Gln Arg His Arg Val Arg Leu
                350                 355                 360
Ala Val Gly Asn Gly Leu Arg Pro Ala Ile Trp Glu Glu Phe Thr
                365                 370                 375
Gln Arg Phe Gly Val Pro Gln Ile Gly Glu Phe Tyr Gly Ala Thr
                380                 385                 390
Glu Cys Asn Cys Ser Ile Ala Asn Met Asp Gly Lys Val Gly Ser
                395                 400                 405
Cys Gly Phe Asn Ser Arg Ile Leu Thr His Val Tyr Pro Ile Arg
                410                 415                 420
Leu Val Lys Val Asn Glu Asp Thr Met Glu Pro Leu Arg Asp Ser
                425                 430                 435
Glu Gly Leu Cys Ile Pro Cys Gln Pro Gly Glu Pro Gly Leu Leu
                440                 445                 450
Val Gly Gln Ile Asn Gln Gln Asp Pro Leu Arg Arg Phe Asp Gly
                455                 460                 465
Tyr Val Ser Asp Ser Ala Thr Asn Lys Lys Ile Ala His Ser Val
                470                 475                 480
Phe Arg Lys Gly Asp Ser Ala Tyr Leu Ser Gly Asp Val Leu Val
                485                 490                 495
Met Asp Glu Leu Gly Tyr Met Tyr Phe Arg Asp Arg Ser Gly Asp
                500                 505                 510
Thr Phe Arg Trp Arg Gly Glu Asn Val Ser Thr Thr Glu Val Glu
```

515 520 525
Ala Val Leu Ser Arg Leu Leu Gly Gln Thr Asp Val Ala Val Tyr
530 535 540
Gly Val Ala Val Pro Gly Val Glu Gly Lys Ala Gly Met Ala Ala
545 550 555
Ile Ala Asp Pro His Ser Gln Leu Asp Pro Asn Ser Met Tyr Gln
560 565 570
Glu Leu Gln Lys Val Leu Ala Ser Tyr Ala Arg Pro Ile Phe Leu
575 580 585
Arg Leu Leu Pro Gln Val Asp Thr Thr Gly Thr Phe Lys Ile Gln
590 595 600
Lys Thr Arg Leu Gln Arg Glu Gly Phe Asp Pro Arg Gln Thr Ser
605 610 615
Asp Arg Leu Phe Phe Leu Asp Leu Lys Gln Gly Arg Tyr Val Pro
620 625 630
Leu Asp Glu Arg Val His Ala Arg Ile Cys Ala Gly Asp Phe Ser
635 640 645
Leu

<210> 36
<211> 691
<212> PRT
<213> Schistosoma mansoni

<300>
<308> GenBank ID No: g425474

<400> 36
Met Phe Ser Ala Leu Cys Arg Arg Gly Phe Leu Thr Asn Lys Val
1 5 10 15
Ser Gln Phe Arg Ser Thr Tyr Lys Cys Asp His Tyr Asn Leu Lys
20 25 30
Thr His Ile Lys Pro Leu Lys Cys Ser Ser Ser Leu Arg Leu Thr
35 40 45
Val Gly Thr Gly Leu Phe Ile Ala Leu His Ser Lys Ile Ser Pro
50 55 60
Glu Ser Arg Ile Gln Thr Val Gln Cys Glu Val Asp Ser Tyr Gln
65 70 75
Thr Asp Gln Ile Thr Phe Ala Lys Ser Gly Gly Ile Pro Arg Tyr
80 85 90
Ile Gly Val Leu Ile Leu Pro Asp Cys Val Tyr Leu Phe Gly Ala
95 100 105
Ile Leu Gly Ala Phe Val Ala Ala Val Met Asn Val Tyr Ile Pro
110 115 120
Leu Tyr Leu Gly Asp Phe Val Ser Ser Leu Ser Arg Cys Val Val
125 130 135
Thr His Glu Gly Phe Val Ser Ala Val Tyr Val Pro Thr Leu Arg
140 145 150
Leu Cys Ser Ser Tyr Leu Leu Gln Ser Leu Ser Thr Phe Leu Tyr
155 160 165
Ile Gly Leu Leu Gly Ser Val Gly Glu Arg Met Ala Arg Arg Met
170 175 180
Arg Ile Gln Leu Phe Arg Lys Leu Val Tyr Gln Asp Val Ala Tyr
185 190 195

```
Phe Asp Val His Ser Ser Gly Lys Leu Val Glu Ile Ile Gly Ser
                200                 205                 210
Asp Val Gln Asn Phe Lys Ser Ser Phe Lys Gln Cys Ile Ser Gln
                215                 220                 225
Gly Leu Arg Asn Gly Ile Gln Val Val Gly Ser Val Phe Ala Leu
                230                 235                 240
Leu Ser Ile Ser Pro Thr Leu Thr Ala Ala Leu Ile Gly Cys Leu
                245                 250                 255
Pro Cys Val Phe Leu Ile Gly Ser Leu Met Gly Thr Glu Leu Arg
                260                 265                 270
His Ile Ser Arg Glu Val Gln Ser Gln Asn Ser Leu Phe Ala Ser
                275                 280                 285
Leu Ile Asp Glu Ala Phe Ser His Ile Arg Thr Val Lys Ser Leu
                290                 295                 300
Ala Met Glu Asp Phe Leu Ile Asn Lys Ile Asn Tyr Asn Val Asp
                305                 310                 315
Lys Ala Lys Met Leu Ser Glu Lys Leu Ser Phe Gly Ile Gly Ser
                320                 325                 330
Phe Gln Gly Leu Ser Asn Leu Thr Leu Asn Gly Val Val Leu Gly
                335                 340                 345
Val Leu Tyr Val Gly Gly His Leu Met Ser Arg Gly Glu Leu Asp
                350                 355                 360
Ala Gly His Leu Met Ser Phe Leu Ala Thr Thr Gln Thr Leu Gln
                365                 370                 375
Arg Ser Leu Thr Gln Leu Ser Leu Leu Tyr Gly Gln Val Val Arg
                380                 385                 390
Gly Tyr Thr Ala Leu Lys Arg Ile His Asp Ile Leu Ala Leu Pro
                395                 400                 405
Ser Gly Ile Gly Ser Ile Pro Ser Ser Ser Ser Ser Leu Val Val
                410                 415                 420
Ser Lys Gln His Val Asn Asn Ile Lys Glu Leu Pro Ser Ser Ser
                425                 430                 435
Ile Tyr Ser Ala Pro Ser Ile Glu Phe Ser Asp Val Lys Phe Ala
                440                 445                 450
Tyr Pro Asn Arg Pro Glu Thr Ile Val Leu Asn Glu Leu Ser Met
                455                 460                 465
Phe Leu Pro Gly Gly Lys Val Ile Ala Leu Val Gly Gln Ser Gly
                470                 475                 480
Ala Gly Lys Ser Thr Val Val Ser Leu Leu Glu Arg Phe Tyr Asp
                485                 490                 495
Pro Ile Ser Gly Glu Ile Leu Leu Asn Gly Asp Lys Leu Thr Asn
                500                 505                 510
Phe Asn Val Asn Tyr Leu Arg Ser Lys Leu Ile Gly Tyr Ile Ser
                515                 520                 525
Gln Glu Pro Gln Ile Phe Asn Ala Ser Ile Arg Glu Asn Ile Arg
                530                 535                 540
Phe Gly Arg Phe Asp Ala Thr Asp Glu Glu Val Glu Glu Ala Ala
                545                 550                 555
Lys Leu Ala Tyr Ala His Asp Phe Ile Ser Asn Asp Leu Pro Tyr
                560                 565                 570
Gly Tyr Asp Thr Leu Val Gly Gln Gly Thr Gly Thr Ile Ala Gly
                575                 580                 585
Leu Ser Gly Gly Gln Arg Gln Arg Ile Ala Ile Ala Arg Ile Leu
                590                 595                 600
Leu Lys Asn Ala Pro Ile Leu Leu Met Asp Glu Ala Thr Ser Ala
```

605 610 615
Leu Asp Thr Glu Ser Glu Ala Lys Val Gln Asn Ala Leu Asn Asn
620 625 630
Ala Met Lys Gly Arg Thr Val Leu Ile Ile Ala His Arg Leu Ser
635 640 645
Thr Val Arg Lys Ala Asp Leu Ile Leu Val Met Ser Lys Gly Gln
650 655 660
Ile Val Glu Lys Gly Thr His Ser Glu Leu Met Ala Asn His Gly
665 670 675
Tyr Tyr Tyr Asn Leu Val Gln Arg Gln Glu Gly Cys Asp Val Phe
680 685 690
Asp

<210> 37
<211> 634
<212> PRT
<213> Rattus norvegicus

<300>
<308> GenBank ID No: g3015617

<400> 37
Met Thr Val Ala Ser Thr Ala Ala Pro Ser Tyr Thr Thr Ser Asp
1 5 10 15
Thr Asn Arg Val Ile Ser Thr Phe Ser Val Val Asp Tyr Val Val
20 25 30
Phe Gly Leu Leu Leu Val Leu Ser Leu Val Ile Gly Leu Tyr His
35 40 45
Ala Cys Arg Gly Trp Gly Arg His Thr Val Gly Glu Leu Leu Met
50 55 60
Ala Asp Arg Lys Met Gly Cys Leu Pro Val Ala Leu Ser Leu Leu
65 70 75
Ala Thr Phe Gln Ser Ala Val Ala Ile Leu Gly Gly Pro Ala Glu
80 85 90
Ile Tyr Arg Phe Gly Thr Gln Tyr Trp Phe Leu Gly Cys Ser Tyr
95 100 105
Phe Leu Gly Leu Leu Ile Pro Ala His Ile Phe Ile Pro Val Phe
110 115 120
Tyr Arg Leu His Leu Thr Ser Ala Tyr Glu Tyr Leu Glu Leu Arg
125 130 135
Phe Asn Lys Ala Val Arg Ile Cys Gly Thr Val Thr Phe Ile Phe
140 145 150
Gln Met Val Val Tyr Met Gly Val Ala Leu Tyr Ala Pro Ser Leu
155 160 165
Ala Leu Asn Ala Val Thr Gly Phe Asp Leu Trp Leu Ser Val Leu
170 175 180
Ala Leu Gly Ile Val Cys Asn Ile Tyr Thr Ala Leu Gly Gly Leu
185 190 195
Lys Ala Val Ile Trp Thr Asp Val Phe Gln Thr Leu Ile Met Phe
200 205 210
Leu Gly Gln Leu Val Val Ile Ile Val Gly Ala Ala Lys Val Gly
215 220 225
Gly Leu Gly His Val Trp Ala Val Ala Ser Gln His Gly Leu Ile
230 235 240

Ser Gly Ile Glu Leu Asp Pro Asp Pro Phe Val Arg His Thr Phe
245 250 255
Trp Thr Leu Ala Phe Gly Gly Val Phe Met Met Leu Ser Leu Tyr
260 265 270
Gly Val Asn Gln Ala Gln Val Gln Arg Tyr Leu Ser Ser His Ser
275 280 285
Glu Lys Ala Ala Val Leu Ser Cys Tyr Ala Val Phe Pro Cys Gln
290 295 300
Gln Val Ala Leu Cys Met Ser Cys Leu Ile Gly Leu Val Met Phe
305 310 315
Ala Tyr Tyr Lys Lys Tyr Ser Met Ser Pro Gln Gln Glu Gln Ala
320 325 330
Ala Pro Asp Gln Leu Val Leu Tyr Phe Val Met Asp Leu Leu Lys
335 340 345
Asp Met Pro Gly Leu Pro Gly Leu Phe Val Ala Cys Leu Phe Ser
350 355 360
Gly Ser Leu Ser Thr Ile Ser Ser Ala Phe Asn Ser Leu Ala Thr
365 370 375
Val Thr Met Glu Asp Leu Ile Gln Pro Trp Phe Pro Gln Leu Thr
380 385 390
Glu Thr Arg Ala Ile Met Leu Ser Arg Ser Leu Ala Phe Ala Tyr
395 400 405
Gly Leu Val Cys Leu Gly Met Ala Tyr Val Ser Ser His Leu Gly
410 415 420
Ser Val Leu Gln Ala Ala Leu Ser Ile Phe Gly Met Val Gly Gly
425 430 435
Pro Leu Leu Gly Leu Phe Cys Leu Gly Met Phe Phe Pro Cys Ala
440 445 450
Asn Pro Leu Gly Ala Ile Val Gly Leu Leu Thr Gly Leu Thr Met
455 460 465
Ala Phe Trp Ile Gly Ile Gly Ser Ile Val Ser Arg Met Ser Ser
470 475 480
Ala Ala Ala Ser Pro Pro Leu Asn Gly Ser Ser Ser Phe Leu Pro
485 490 495
Ser Asn Leu Thr Val Ala Thr Val Thr Thr Leu Met Pro Ser Thr
500 505 510
Leu Ser Lys Pro Thr Gly Leu Gln Gln Phe Tyr Ser Leu Ser Tyr
515 520 525
Leu Trp Tyr Ser Ala His Asn Ser Thr Thr Val Ile Ala Val Gly
530 535 540
Leu Ile Val Ser Leu Leu Thr Gly Gly Met Arg Gly Arg Ser Leu
545 550 555
Asn Pro Gly Thr Ile Tyr Pro Val Leu Pro Lys Leu Leu Ala Leu
560 565 570
Leu Pro Leu Ser Cys Gln Lys Arg Leu Cys Trp Arg Ser His Asn
575 580 585
Gln Asp Ile Pro Val Val Thr Asn Leu Phe Pro Glu Lys Met Gly
590 595 600
Asn Gly Ala Leu Gln Asp Ser Arg Asp Lys Glu Arg Met Ala Glu
605 610 615
Asp Gly Leu Val His Gln Pro Cys Ser Pro Thr Tyr Ile Val Gln
620 625 630
Glu Thr Ser Leu

<210> 38
<211> 507
<212> PRT
<213> Homo sapiens

<300>
<308> GenBank ID No: g3639058

<400> 38
Met Ala Gly Ala Gly Pro Lys Arg Arg Ala Leu Ala Ala Pro Ala
1 5 10 15
Ala Glu Glu Lys Glu Glu Ala Arg Glu Lys Met Leu Ala Ala Lys
20 25 30
Ser Ala Asp Gly Ser Ala Pro Ala Gly Glu Gly Gly Val Thr
35 40 45
Leu Gln Arg Asn Ile Thr Leu Leu Asn Gly Val Ala Ile Ile Val
50 55 60
Gly Thr Ile Ile Gly Ser Gly Ile Phe Val Thr Pro Thr Gly Val
65 70 75
Leu Lys Glu Ala Gly Ser Pro Gly Leu Ala Leu Val Val Trp Ala
80 85 90
Ala Cys Gly Val Phe Ser Ile Val Gly Ala Leu Cys Tyr Ala Glu
95 100 105
Leu Gly Thr Thr Ile Ser Lys Ser Gly Gly Asp Tyr Ala Tyr Met
110 115 120
Leu Glu Val Tyr Gly Ser Leu Pro Ala Phe Leu Lys Leu Trp Ile
125 130 135
Glu Leu Leu Ile Ile Arg Pro Ser Ser Gln Tyr Ile Val Ala Leu
140 145 150
Val Phe Ala Thr Tyr Leu Leu Lys Pro Leu Phe Pro Thr Cys Pro
155 160 165
Val Pro Glu Glu Ala Ala Lys Leu Val Ala Cys Leu Cys Val Leu
170 175 180
Leu Leu Thr Ala Val Asn Cys Tyr Ser Val Lys Ala Ala Thr Arg
185 190 195
Val Gln Asp Ala Phe Ala Ala Ala Lys Leu Leu Ala Leu Ala Leu
200 205 210
Ile Ile Leu Leu Gly Phe Val Gln Ile Gly Lys Gly Asp Val Ser
215 220 225
Asn Leu Asp Pro Lys Phe Ser Phe Glu Gly Thr Lys Leu Asp Val
230 235 240
Gly Asn Ile Val Leu Ala Leu Tyr Ser Gly Leu Phe Ala Tyr Gly
245 250 255
Gly Trp Asn Tyr Leu Asn Phe Val Thr Glu Glu Met Ile Asn Pro
260 265 270
Tyr Arg Asn Leu Pro Leu Ala Ile Ile Ile Ser Leu Pro Ile Val
275 280 285
Thr Leu Val Tyr Val Leu Thr Asn Leu Ala Tyr Phe Thr Thr Leu
290 295 300
Ser Thr Glu Gln Met Leu Ser Ser Glu Ala Val Ala Val Asp Phe
305 310 315
Gly Asn Tyr His Leu Gly Val Met Ser Trp Ile Ile Pro Val Phe
320 325 330
Val Gly Leu Ser Cys Phe Gly Ser Val Asn Gly Ser Leu Phe Thr

```
                335                 340                 345
Ser Ser Arg Leu Phe Phe Val Gly Ser Arg Glu Gly His Leu Pro
                350                 355                 360
Ser Ile Leu Ser Met Ile His Pro Gln Leu Leu Thr Pro Val Pro
                365                 370                 375
Ser Leu Val Phe Thr Cys Val Met Thr Leu Leu Tyr Ala Phe Ser
                380                 385                 390
Lys Asp Ile Phe Ser Val Ile Asn Phe Phe Ser Phe Phe Asn Trp
                395                 400                 405
Leu Cys Val Ala Leu Ala Ile Ile Gly Met Ile Trp Leu Arg His
                410                 415                 420
Arg Lys Pro Glu Leu Glu Arg Pro Ile Lys Val Asn Leu Ala Leu
                425                 430                 435
Pro Val Phe Phe Ile Leu Ala Cys Leu Phe Leu Ile Ala Val Ser
                440                 445                 450
Phe Trp Lys Thr Pro Val Glu Cys Gly Ile Gly Phe Thr Ile Ile
                455                 460                 465
Leu Ser Gly Leu Pro Val Tyr Phe Phe Gly Val Trp Trp Lys Asn
                470                 475                 480
Lys Pro Lys Trp Leu Leu Gln Gly Ile Phe Ser Thr Thr Val Leu
                485                 490                 495
Cys Gln Lys Leu Met Gln Val Val Pro Gln Glu Thr
                500                 505
```

<210> 39
<211> 504
<212> PRT
<213> Homo sapiens

<300>
<308> GenBank ID No: g1840045

<400> 39

```
Met Glu Ala Pro Leu Gln Thr Glu Met Val Glu Leu Val Pro Asn
  1               5                  10                  15
Gly Lys His Ser Glu Gly Leu Leu Pro Val Ile Thr Pro Met Ala
                20                  25                  30
Gly Asn Gln Arg Val Glu Asp Pro Ala Arg Ser Cys Met Glu Gly
                35                  40                  45
Lys Ser Phe Leu Gln Lys Ser Pro Ser Lys Glu Pro His Phe Thr
                50                  55                  60
Asp Phe Glu Gly Lys Thr Ser Phe Gly Met Ser Val Phe Asn Leu
                65                  70                  75
Ser Asn Ala Ile Met Gly Ser Gly Ile Leu Gly Leu Ala Tyr Ala
                80                  85                  90
Met Ala Asn Thr Gly Ile Ile Leu Phe Leu Phe Leu Leu Thr Ala
                95                 100                 105
Val Ala Leu Leu Ser Ser Tyr Ser Ile His Leu Leu Leu Lys Ser
               110                 115                 120
Ser Gly Val Val Gly Ile Arg Ala Tyr Glu Gln Leu Gly Tyr Arg
               125                 130                 135
Ala Phe Gly Thr Pro Gly Lys Leu Ala Ala Ala Leu Ala Ile Thr
               140                 145                 150
```

```
Leu Gln Asn Ile Gly Ala Met Ser Ser Tyr Leu Tyr Ile Ile Lys
                155                 160                 165
Ser Glu Leu Pro Leu Val Ile Gln Thr Phe Leu Asn Leu Glu Glu
                170                 175                 180
Lys Thr Ser Asp Trp Tyr Met Asn Gly Asn Tyr Leu Val Ile Leu
                185                 190                 195
Val Ser Val Thr Ile Ile Leu Pro Leu Ala Leu Met Arg Gln Leu
                200                 205                 210
Gly Tyr Leu Gly Tyr Ser Ser Gly Phe Ser Leu Ser Cys Met Val
                215                 220                 225
Phe Phe Leu Ile Ala Val Ile Tyr Lys Lys Phe His Val Pro Cys
                230                 235                 240
Pro Leu Pro Pro Asn Phe Asn Asn Thr Thr Gly Asn Phe Ser His
                245                 250                 255
Val Glu Ile Val Lys Glu Lys Val Gln Leu Gln Val Glu Pro Glu
                260                 265                 270
Ala Ser Ala Phe Cys Thr Pro Ser Tyr Phe Thr Leu Asn Ser Gln
                275                 280                 285
Thr Ala Tyr Thr Ile Pro Ile Met Ala Phe Ala Phe Val Cys His
                290                 295                 300
Pro Glu Val Leu Pro Ile Tyr Thr Glu Leu Lys Asp Pro Ser Lys
                305                 310                 315
Lys Lys Met Gln His Ile Ser Asn Leu Ser Ile Ala Val Met Tyr
                320                 325                 330
Ile Met Tyr Phe Leu Ala Ala Leu Phe Gly Tyr Leu Thr Phe Tyr
                335                 340                 345
Asn Gly Val Glu Ser Glu Leu Leu His Thr Tyr Ser Lys Val Asp
                350                 355                 360
Pro Phe Asp Val Leu Ile Leu Cys Val Arg Val Ala Val Leu Thr
                365                 370                 375
Ala Val Thr Leu Thr Val Pro Ile Val Leu Phe Pro Val Arg Arg
                380                 385                 390
Ala Ile Gln Gln Met Leu Phe Pro Asn Gln Glu Phe Ser Trp Leu
                395                 400                 405
Arg His Val Leu Ile Ala Val Gly Leu Leu Thr Cys Ile Asn Leu
                410                 415                 420
Leu Val Ile Phe Ala Pro Asn Ile Leu Gly Ile Phe Gly Val Ile
                425                 430                 435
Gly Ala Thr Ser Ala Pro Phe Leu Ile Phe Ile Phe Pro Ala Ile
                440                 445                 450
Phe Tyr Phe Arg Ile Met Pro Thr Glu Lys Glu Pro Ala Arg Ser
                455                 460                 465
Thr Pro Lys Ile Leu Ala Leu Cys Phe Ala Met Leu Gly Phe Leu
                470                 475                 480
Leu Met Thr Met Ser Leu Ser Phe Ile Ile Ile Asp Trp Ala Ser
                485                 490                 495
Gly Thr Ser Arg His Gly Gly Asn His
                500
```

<210> 40
<211> 393
<212> PRT
<213> Homo sapiens

<300>
<308> GenBank ID No: g1526438

<400> 40
Met Ala Ala Val Gly Ala Gly Gly Ser Thr Ala Ala Pro Gly Pro
1 5 10 15
Gly Ala Val Ser Ala Gly Ala Leu Glu Pro Gly Thr Ala Ser Ala
20 25 30
Ala His Arg Arg Leu Lys Tyr Ile Ser Leu Ala Val Leu Val Val
35 40 45
Gln Asn Ala Ser Leu Ile Leu Ser Ile Arg Tyr Ala Arg Thr Leu
50 55 60
Pro Gly Asp Arg Phe Phe Ala Thr Thr Ala Val Val Met Ala Glu
65 70 75
Val Leu Lys Gly Leu Thr Cys Leu Leu Leu Leu Phe Ala Gln Lys
80 85 90
Arg Gly Asn Val Lys His Leu Val Leu Phe Leu His Glu Ala Val
95 100 105
Leu Val Gln Tyr Val Asp Thr Leu Lys Leu Ala Val Pro Ser Leu
110 115 120
Ile Tyr Thr Leu Gln Asn Asn Leu Gln Tyr Val Ala Ile Ser Asn
125 130 135
Leu Pro Ala Ala Thr Phe Gln Val Thr Tyr Gln Leu Lys Ile Leu
140 145 150
Thr Thr Ala Leu Phe Ser Val Leu Met Leu Asn Arg Ser Leu Ser
155 160 165
Arg Leu Gln Trp Ala Ser Leu Leu Leu Leu Phe Thr Gly Val Ala
170 175 180
Ile Val Gln Ala Gln Gln Ala Gly Gly Gly Gly Pro Arg Pro Leu
185 190 195
Asp Gln Asn Pro Gly Ala Gly Leu Ala Ala Val Val Ala Ser Cys
200 205 210
Leu Ser Ser Gly Phe Ala Gly Val Tyr Phe Glu Lys Ile Leu Lys
215 220 225
Gly Ser Ser Gly Ser Val Trp Leu Arg Asn Leu Gln Leu Gly Leu
230 235 240
Phe Gly Thr Ala Leu Gly Leu Val Gly Leu Trp Trp Ala Glu Gly
245 250 255
Thr Ala Val Ala Thr Arg Gly Phe Phe Phe Gly Tyr Thr Pro Ala
260 265 270
Val Trp Gly Val Val Leu Asn Gln Ala Phe Gly Gly Leu Leu Val
275 280 285
Ala Val Val Val Lys Tyr Ala Asp Asn Ile Leu Lys Gly Phe Ala
290 295 300
Thr Ser Leu Ser Ile Val Leu Ser Thr Val Ala Ser Ile Arg Leu
305 310 315
Phe Gly Phe His Val Asp Pro Leu Phe Ala Leu Gly Ala Gly Leu
320 325 330
Val Ile Gly Ala Val Tyr Leu Tyr Ser Leu Pro Arg Gly Ala Ala
335 340 345
Lys Ala Ile Ala Ser Ala Ser Ala Ser Ala Ser Gly Pro Cys Val
350 355 360
His Gln Gln Pro Pro Gly Gln Pro Pro Pro Pro Gln Leu Ser Ser
365 370 375
His Arg Gly Asp Leu Ile Thr Glu Pro Phe Leu Pro Lys Ser Val

380 385 390
Leu Val Lys

<210> 41
<211> 893
<212> PRT
<213> Homo sapiens

<300>
<308> GenBank ID No: g3335175

<400> 41
His Val Gln Asp Phe Thr Ala Phe Trp Asp Lys Ala Ser Glu Thr
1 5 10 15
Pro Thr Leu Gln Gly Leu Ser Phe Thr Val Arg Pro Gly Glu Leu
20 25 30
Leu Ala Val Val Gly Pro Val Gly Ala Gly Lys Ser Ser Leu Leu
35 40 45
Ser Ala Val Leu Gly Glu Leu Ala Pro Ser His Gly Leu Val Ser
50 55 60
Val His Gly Arg Ile Ala Tyr Val Ser Gln Gln Pro Trp Val Phe
65 70 75
Ser Gly Thr Leu Arg Ser Asn Ile Leu Phe Gly Lys Lys Tyr Glu
80 85 90
Lys Glu Arg Tyr Glu Lys Val Ile Lys Ala Cys Ala Leu Lys Lys
95 100 105
Asp Leu Gln Leu Leu Glu Asp Gly Asp Leu Thr Val Ile Gly Asp
110 115 120
Arg Gly Thr Thr Leu Ser Gly Gly Gln Lys Ala Arg Val Asn Leu
125 130 135
Ala Arg Ala Val Tyr Gln Asp Ala Asp Ile Tyr Leu Leu Asp Asp
140 145 150
Pro Leu Ser Ala Val Asp Ala Glu Val Ser Arg His Leu Phe Glu
155 160 165
Leu Cys Ile Cys Gln Ile Leu His Glu Lys Ile Thr Ile Leu Val
170 175 180
Thr His Gln Leu Gln Tyr Leu Lys Ala Ala Ser Gln Ile Leu Ile
185 190 195
Leu Lys Asp Gly Lys Met Val Gln Lys Gly Thr Tyr Thr Glu Phe
200 205 210
Leu Lys Ser Gly Ile Asp Phe Gly Ser Leu Leu Lys Lys Asp Asn
215 220 225
Glu Glu Ser Glu Gln Pro Pro Val Pro Gly Thr Pro Thr Leu Arg
230 235 240
Asn Arg Thr Phe Ser Glu Ser Ser Val Trp Ser Gln Gln Ser Ser
245 250 255
Arg Pro Ser Leu Lys Asp Gly Ala Leu Glu Ser Gln Asp Thr Glu
260 265 270
Asn Val Pro Val Thr Leu Ser Glu Glu Asn Arg Ser Glu Gly Lys
275 280 285
Val Gly Phe Gln Ala Tyr Lys Asn Tyr Phe Arg Ala Gly Ala His
290 295 300
Trp Ile Val Phe Ile Phe Leu Ile Leu Leu Asn Thr Ala Ala Gln

305 310 315
Val Ala Tyr Val Leu Gln Asp Trp Trp Leu Ser Tyr Trp Ala Asn
320 325 330
Lys Gln Ser Met Leu Asn Val Thr Val Asn Gly Gly Gly Asn Val
335 340 345
Thr Glu Lys Leu Asp Leu Asn Trp Tyr Leu Gly Ile Tyr Ser Gly
350 355 360
Leu Thr Val Ala Thr Val Leu Phe Gly Ile Ala Arg Ser Leu Leu
365 370 375
Val Phe Tyr Val Leu Val Asn Ser Ser Gln Thr Leu His Asn Lys
380 385 390
Met Phe Glu Ser Ile Leu Lys Ala Pro Val Leu Phe Phe Asp Arg
395 400 405
Asn Pro Ile Gly Arg Ile Leu Asn Arg Phe Ser Lys Asp Ile Gly
410 415 420
His Leu Asp Asp Leu Leu Pro Leu Thr Phe Leu Asp Phe Ile Gln
425 430 435
Thr Leu Leu Gln Val Val Gly Val Val Ser Val Ala Val Ala Val
440 445 450
Ile Pro Trp Ile Ala Ile Pro Leu Val Pro Leu Gly Ile Ile Phe
455 460 465
Ile Phe Leu Arg Arg Tyr Phe Leu Glu Thr Ser Arg Asp Val Lys
470 475 480
Arg Leu Glu Ser Thr Thr Arg Ser Pro Val Phe Ser His Leu Ser
485 490 495
Ser Ser Leu Gln Gly Leu Trp Thr Ile Arg Ala Tyr Lys Ala Glu
500 505 510
Glu Arg Cys Gln Glu Leu Phe Asp Ala His Gln Asp Leu His Ser
515 520 525
Glu Ala Trp Phe Leu Phe Leu Thr Thr Ser Arg Trp Phe Ala Val
530 535 540
Arg Leu Asp Ala Ile Cys Ala Met Phe Val Ile Ile Val Ala Phe
545 550 555
Gly Ser Leu Ile Leu Ala Lys Thr Leu Asp Ala Gly Gln Val Gly
560 565 570
Leu Ala Leu Ser Tyr Ala Leu Thr Leu Met Gly Met Phe Gln Trp
575 580 585
Cys Val Arg Gln Ser Ala Glu Val Glu Asn Met Met Ile Ser Val
590 595 600
Glu Arg Val Ile Glu Tyr Thr Asp Leu Glu Lys Glu Ala Pro Trp
605 610 615
Glu Tyr Gln Lys Arg Pro Pro Pro Ala Trp Pro His Glu Gly Val
620 625 630
Ile Ile Phe Asp Asn Val Asn Phe Met Tyr Ser Pro Gly Gly Pro
635 640 645
Leu Val Leu Lys His Leu Thr Ala Leu Ile Lys Ser Gln Glu Lys
650 655 660
Val Gly Ile Val Gly Arg Thr Gly Ala Gly Lys Ser Ser Leu Ile
665 670 675
Ser Ala Leu Phe Arg Leu Ser Glu Pro Glu Gly Lys Ile Trp Ile
680 685 690
Asp Lys Ile Leu Thr Thr Glu Ile Gly Leu His Asp Leu Arg Lys
695 700 705
Lys Met Ser Ile Ile Pro Gln Glu Pro Val Leu Phe Thr Gly Thr
710 715 720

```
Met Arg Lys Asn Leu Asp Pro Phe Lys Glu His Thr Asp Glu Glu
                725                 730                 735
Leu Trp Asn Ala Leu Gln Glu Val Gln Leu Lys Glu Thr Ile Glu
                740                 745                 750
Asp Leu Pro Gly Lys Met Asp Thr Glu Leu Ala Glu Ser Gly Ser
                755                 760                 765
Asn Phe Ser Val Gly Gln Arg Gln Leu Val Cys Leu Ala Arg Ala
                770                 775                 780
Ile Leu Arg Lys Asn Gln Ile Leu Ile Ile Asp Glu Ala Thr Ala
                785                 790                 795
Asn Val Asp Pro Arg Thr Asp Glu Leu Ile Gln Lys Lys Ile Arg
                800                 805                 810
Glu Lys Phe Ala His Cys Thr Val Leu Thr Ile Ala His Arg Leu
                815                 820                 825
Asn Thr Ile Ile Asp Ser Asp Lys Ile Met Val Leu Asp Ser Gly
                830                 835                 840
Arg Leu Lys Glu Tyr Asp Glu Pro Tyr Val Leu Leu Gln Asn Lys
                845                 850                 855
Glu Ser Leu Phe Tyr Lys Met Val Gln Gln Leu Gly Lys Ala Glu
                860                 865                 870
Ala Ala Ala Leu Thr Glu Thr Ala Lys Gln Val Ile Leu Gln Lys
                875                 880                 885
Lys Leu Ser Thr Tyr Trp Ser His
                890
```

<210> 42
<211> 453
<212> PRT
<213> Homo sapiens

<300>
<308> GenBank ID No: g1279457

<400> 42

```
Met Ala Leu Arg Gly Phe Cys Ser Arg Trp Leu Arg Pro Ala Leu
  1                 5                  10                  15
Ala Ile Gly Leu Phe Ala Ser Met Ala Ala Val Leu Leu Gly Gly
                 20                  25                  30
Ala Arg Ala Ser Arg Leu Leu Phe Gln Arg Leu Leu Trp Asp Val
                 35                  40                  45
Val Arg Ser Pro Ile Ser Phe Phe Glu Arg Thr Pro Ile Gly His
                 50                  55                  60
Leu Leu Asn Arg Phe Ser Lys Glu Thr Asp Thr Val Asp Val Asp
                 65                  70                  75
Ile Pro Asp Lys Leu Arg Ser Leu Leu Met Tyr Ala Phe Gly Leu
                 80                  85                  90
Leu Glu Val Ser Leu Val Val Glu Trp Pro Thr Pro Leu Pro Leu
                 95                 100                 105
Trp Pro Ser Cys His Cys Phe Ser Ser Thr Leu Gly Phe Arg Trp
                110                 115                 120
Leu Ala Ala Asn Val Glu Leu Leu Gly Asn Gly Leu Val Phe Ala
                125                 130                 135
Ala Ala Thr Cys Ala Val Leu Ser Lys Ala His Leu Ser Ala Gly
                140                 145                 150
```

```
Leu Val Gly Phe Ser Val Ser Ala Ala Leu Gln Val Thr Gln Thr
                155                 160                 165
Leu Gln Trp Val Val Arg Asn Trp Thr Asp Leu Glu Asn Ser Ile
                170                 175                 180
Val Ser Val Glu Arg Met Gln Asp Tyr Ala Trp Thr Pro Lys Glu
                185                 190                 195
Ala Pro Trp Arg Leu Pro Thr Cys Ala Ala Gln Pro Pro Trp Pro
                200                 205                 210
Gln Gly Gly Gln Ile Glu Phe Arg Asp Phe Gly Leu Arg Tyr Arg
                215                 220                 225
Pro Glu Leu Pro Leu Ala Val Gln Gly Val Ser Phe Lys Ile His
                230                 235                 240
Ala Gly Glu Lys Val Gly Ile Val Gly Arg Thr Gly Ala Gly Lys
                245                 250                 255
Ser Ser Leu Ala Ser Gly Leu Leu Arg Leu Gln Glu Ala Ala Glu
                260                 265                 270
Gly Gly Ile Trp Ile Asp Gly Val Pro Ile Ala His Val Gly Val
                275                 280                 285
His Thr Leu Arg Ser Arg Ile Ser Ile Ile Pro Gln Asp Pro Ile
                290                 295                 300
Leu Phe Pro Gly Ser Leu Arg Met Asn Leu Asp Leu Leu Gln Glu
                305                 310                 315
His Ser Asp Glu Ala Ile Trp Ala Ala Leu Glu Thr Val Gln Leu
                320                 325                 330
Lys Ala Leu Val Ala Cys Leu Pro Gly Gln Leu Gln Tyr Lys Cys
                335                 340                 345
Ala Asp Arg Gly Glu Asp Leu Ser Val Gly Gln Lys Gln Leu Leu
                350                 355                 360
Cys Leu Ala Arg Ala Leu Leu Arg Lys Thr Gln Ile Leu Ile Leu
                365                 370                 375
Asp Glu Ala Thr Ala Ala Val Asp Pro Gly Thr Glu Leu Gln Met
                380                 385                 390
Gln Ala Met Leu Gly Ser Trp Phe Ala Gln Cys Thr Val Leu Leu
                395                 400                 405
Ile Ala His Arg Leu Arg Ser Val Met Asp Cys Ala Arg Val Leu
                410                 415                 420
Val Met Asp Lys Gly Gln Val Ala Glu Ser Gly Ser Pro Ala Gln
                425                 430                 435
Leu Leu Ala Gln Lys Gly Leu Phe Tyr Arg Leu Ala Gln Glu Ser
                440                 445                 450
Gly Leu Val
```

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
11 May 2000 (11.05.2000)

PCT

(10) International Publication Number
WO 00/26245 A3

(51) International Patent Classification[7]: C12N 15/12, C07K 14/47, 16/18, A61K 38/17, C12Q 1/68, C12N 1/21

(21) International Application Number: PCT/US99/26048

(22) International Filing Date:
4 November 1999 (04.11.1999)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/172,255	4 November 1998 (04.11.1998)	US
60/172,252	24 November 1998 (24.11.1998)	US
60/172,214	22 December 1998 (22.12.1998)	US
60/121,896	26 February 1999 (26.02.1999)	US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	60/172,255 (CIP)
Filed on	4 November 1998 (04.11.1998)
US	60/172,252 (CIP)
Filed on	24 November 1998 (24.11.1998)
US	60/172,214 (CIP)
Filed on	22 December 1998 (22.12.1998)
US	60/121,896 (CIP)
Filed on	26 February 1999 (26.02.1999)

(71) Applicant *(for all designated States except US)*: **INCYTE PHARMACEUTICALS, INC.** [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **HILLMAN, Jennifer, L.** [US/US]; 230 Monroe Drive, #12, Mountain View, CA 94040 (US). **YUE, Henry** [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). **TANG, Y., Tom** [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **LAL, Preeti** [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). **CORLEY, Neil, C.** [US/US]; 1240 Dale Avenue, #30, Mountain View, CA 94040 (US). **GUEGLER, Karl, J.** [CH/US]; 1048 Oakland Avenue, Menlo Park, CA 94025 (US). **BAUGHN, Mariah, R.** [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). **AZIMZAI, Yalda** [US/US]; 2045 Rock Springs Drive, Hayward, CA 94545 (US). **LU, Dyung, Aina, M.** [US/US]; 55 Park Belmont Place, San Jose, CA 95136 (US).

(74) Agents: **BILLINGS, Lucy, J.** et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States *(national)*: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *With international search report.*

(88) Date of publication of the international search report:
19 April 2001

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 00/26245 A3

(54) Title: HUMAN MEMBRANE TRANSPORT PROTEINS

(57) Abstract: The invention provides human membrane transport proteins (MTRP) and polynucleotides which identify and encode MTRP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of MTRP.

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/47 C07K16/18 A61K38/17 C12Q1/68
C12N1/21

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C07K C12N A61K C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	DATABASE GENEMBL [Online] 20 February 1998 (1998-02-20) HALLECK M.S. ET AL: "Homo sapiens putative ATPase mRNA, partial cds." XP002132693 Accession No. U78978 -& HALLECK, M.S. ET AL.: "Multiple members of a third subfamily of P-type ATPases identified by genomic sequences and ESTs." GENOME RESEARCH, vol. 8, no. 4, April 1998 (1998-04), pages 354-361, XP002132690 figure 2B page 357 --- -/--	1-16,19, 20

[X] Further documents are listed in the continuation of box C. [ ] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not coneidered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
15 March 2000	20. JUNI. 2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	ALCONADA RODRIG.., A

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE GENEMBL [Online] 27 May 1998 (1998-05-27) STANCHI,F.: "Homo sapiens mRNA for putative ATPase, partial" XP002132694 Accession AJ006268 ---	1-16,19, 20
A	ALLIKMETS R ET AL: "Characterization of the human ABC superfamily: isolation and mapping of 21 new genes using the expressed sequence tags database." HUMAN MOLECULAR GENETICS, (1996 OCT) 5 (10) 1649-55., XP002132691 figures 2,3; table 1 ---	
A	MASTROBERARDINO L ET AL: "Amino-acid transport by heterodimers of 4F2hc/CD98 and members of a permease family." NATURE, (1998 SEP 17) 395 (6699) 288-91., XP002132692 the whole document ---	
A	SARDET C ET AL: "Molecular cloning, primary structure, and expression of the human growth factor-activatable Na+/H+ antiporter." CELL, (1989 JAN 27) 56 (2) 271-80., XP000876824 the whole document -----	

INTERNATIONAL SEARCH REPORT

Int. onal application No. PCT/US 99/26048

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claims 19 and 20 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [X] Claims Nos.: 17,18
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [X] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-16,19-20 (partially)

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

1. Claims: 1-16,19-20 (partially)

Human polypeptide comprising the amino acid sequence of SEQ ID NO:1 or any variant having at least 90% amino acid identity with said sequence, the polynucleotide sequence encoding said polypeptide (SEQ ID NO:18) and any variant having at least 90% identity with said polynucleotide sequence; a polynucleotide that hybridizes with said polynucleotide; methods for detection of the polynucleotide; expression vectors and hosts for the recombinant expression of said polypeptide; method for the production of said polypeptide; antibodies against said polypeptide, agonists, antagonists, pharmaceutical compositions containing said polypeptide and uses thereof for the treatment or prevention of a disorder.

2. Claims: 1-16,19-20 (partially)

As subject 1, but comprising the polypeptide sequence of SEQ ID NO:2 and the polynucleotide sequence of SEQ ID NO:19.

3. Claims: 1-16,19-20 (partially)

As subject 1, but comprising the polypeptide sequence of SEQ ID NO:3 and the polynucleotide sequence of SEQ ID NO:20.

4. Claims: 1-16,19-20 (partially)

As subject 1, but comprising the polypeptide sequence of SEQ ID NO:4 and the polynucleotide sequence of SEQ ID NO:21.

5. Claims: 1-16,19-20 (partially)

As subject 1, but comprising the polypeptide sequence of SEQ ID NO:5 and the polynucleotide sequence of SEQ ID NO:22.

6. Claims: 1-16,19-20 (partially)

As subject 1, but comprising the polypeptide sequence of SEQ ID NO:6 and the polynucleotide sequence of SEQ ID NO:23.

7. Claims: 1-16,19-20 (partially)

As subject 1, but comprising the polypeptide sequences of SEQ ID NO:7 and 11 and the respective polynucleotide sequences of SEQ ID NO:24 and 28.

8. Claims: 1-16,19-20 (partially)

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

As subject 1, but comprising the polypeptide sequence of SEQ ID NO:8 and the polynucleotide sequence of SEQ ID NO:25.

9. Claims: 1-16,19-20 (partially)

As subject 1, but comprising the polypeptide sequence of SEQ ID NO:9 and the polynucleotide sequence of SEQ ID NO:26.

10. Claims: 1-16,19-20 (partially)

As subject 1, but comprising the polypeptide sequence of SEQ ID NO:10 and the polynucleotide sequence of SEQ ID NO:27.

11. Claims: 1-16,19-20 (partially)

As subject 1, but comprising the polypeptide sequence of SEQ ID NO:12 and the polynucleotide sequence of SEQ ID NO:29.

12. Claims: 1-16,19-20 (partially)

As subject 1, but comprising the polypeptide sequence of SEQ ID NO:13 and the polynucleotide sequence of SEQ ID NO:30.

13. Claims: 1-16,19-20 (partially)

As subject 1, but comprising the polypeptide sequence of SEQ ID NO:14 and the polynucleotide sequence of SEQ ID NO:31.

14. Claims: 1-16,19-20 (partially)

As subject 1, but comprising the polypeptide sequence of SEQ ID NO:15 and the polynucleotide sequence of SEQ ID NO:32.

15. Claims: 1-16,19-20 (partially)

As subject 1, but comprising the polypeptide sequence of SEQ ID NO:16 and the polynucleotide sequence of SEQ ID NO:33.

16. Claims: 1-16,19-20 (partially)

As subject 1, but comprising the polypeptide sequence of SEQ ID NO:17 and the polynucleotide sequence of SEQ ID NO:34.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 17,18

Present claims 17 and 18 relate to an extremely large number of possible compounds. Support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT is not to be found, however, for any of the compounds claimed. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Consequently, the search has not been carried out for those claims which do not appear to be supported and disclosed, namely those parts relating to the agonists and antagonists of the polypeptides of the invention.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.